

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 19:04:56 ; Search time 52.0548 Seconds
(without alignments)
186.909 Million cell updates/sec

Title: US-09-632-036f-1

Perfect score: 100

Sequence: 1 TGTDMKRLPASPETHLDM 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 46614

Minimum DB seq length: 0
Maximum DB seq length: 40

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Uniprot_03:*

1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	36.0	33	039812	039812 encephalomy
2	36	36.0	33	039814	039814 encephalomy
3	36	36.0	33	039816	039816 encephalomy
4	34	34.0	35	09FGV7	09FGV7 xylella fas
5	32	32.0	18	047137	047137 escherichia
6	32	32.0	20	09R636	09R636 desulfovibr
7	32	32.0	40	07QNU4	07QNU4 giardia lam
8	32	32.0	40	0677P2	0677P2 lymphocysti
9	31	31.0	21	09UC66	09UC66 homo sapien
10	31	31.0	31	044263	044263 aeromonas s
11	31	31.0	38	08FWY9	08FWY9 brucella su
12	31	31.0	39	08QYN4	08QYN4 potato leaf
13	30	30.0	20	07M212	07M212 glycine max
14	30	30.0	30	0878C0	0878C0 streptococc
15	30	30.0	31	09Z516	09Z516 streptomyc
16	30	30.0	33	08MKG5	08MKG5 sus scrofa
17	30	30.0	39	08EB82	08EB82 shewanella
18	29	29.0	16	09UMJ9	09UMJ9 homo sapien
19	29	29.0	19	090RF8	090RF8 human immun
20	29	29.0	26	09RT76	09RT76 sus scrofa
21	29	29.0	35	033825	033825 anas platyr
22	29	29.0	36	08XRW4	08XRW4 raietonia s
23	29	29.0	37	09NFF7	09NFF7 plasmodium
24	29	29.0	38	0811B9	0811B9 mus musculu
25	28	28.0	18	09TWM9	09TWM9 tetrahymena
26	28	28.0	27	023749	023749 ctenodrilus
27	28	28.0	30	06R810	06R810 streptomyc
28	28	28.0	31	09TRI6	09TRI6 sus scrofa
29	28	28.0	33	09UKM0	09UKM0 homo sapien
30	28	28.0	33	09JU69	09JU69 neisseria m
31	28	28.0	36	004726	004726 silene lati

32	28	28.0	40	2	P78341	P78341 homo sapien
33	28	28.0	40	2	P78342	P78342 homo sapien
34	28	28.0	40	2	Q924Y1	Q924Y1 rattus norv
35	27.5	27.5	39	2	Q7UQD8	Q7UQD8 rhodopirell
36	27	27.0	20	2	Q7SBL0	Q7SBL0 neurospora
37	27	27.0	20	2	P79837	P79837 prionalluru
38	27	27.0	26	1	PSBO	PSBO PINPS
39	27	27.0	26	2	Q51935	Q51935 plasmid nri
40	27	27.0	27	2	Q6WBJ7	Q6WBJ7 liolaemus g
41	27	27.0	27	2	Q978S6	Q978S6 liolaemus c
42	27	27.0	31	2	Q9LEJ3	Q9LEJ3 glycine max
43	27	27.0	31	2	Q9KP96	Q9KP96 vibrio chol
44	27	27.0	32	2	Q7SBR6	Q7SBR6 neurospora
45	27	27.0	34	2	Q8C4P4	Q8C4P4 mus musculu
46	27	27.0	34	2	Q6WP73	Q6WP73 cathartes m
47	27	27.0	34	2	Q6WP74	Q6WP74 aegothelies
48	27	27.0	34	2	Q6WP75	Q6WP75 podargus st
49	27	27.0	34	2	Q6WP76	Q6WP76 podager nac
50	27	27.0	34	2	Q6WP77	Q6WP77 nyctibius a
51	27	27.0	34	2	Q6WP78	Q6WP78 accipiter g
52	27	27.0	34	2	Q6WP79	Q6WP79 falco pereg
53	27	27.0	34	2	Q6WP80	Q6WP80 tyto alba (
54	27	27.0	34	2	Q6WP81	Q6WP81 strix urale
55	27	27.0	34	2	Q6WP82	Q6WP82 bubo bubo (
56	27	27.0	34	2	Q6WP83	Q6WP83 sturnus vul
57	27	27.0	34	2	Q6WP84	Q6WP84 anas platyr
58	27	27.0	35	2	Q7QVP9	Q7QVP9 giardia lam
59	27	27.0	36	2	Q74G13	Q74G13 geobacter s
60	27	27.0	37	2	Q56388	Q56388 pseudomonas
61	27	27.0	37	2	Q6QH81	Q6QH81 escherichia
62	27	27.0	38	2	Q6NSE6	Q6NSE6 homo sapien
63	27	27.0	38	2	Q86MJ0	Q86MJ0 caenorhabdi
64	27	27.0	38	2	Q59973	Q59973 shigella so
65	27	27.0	38	2	Q6QH79	Q6QH79 escherichia
66	27	27.0	39	2	Q6IT77	Q6IT77 homo sapien
67	27	27.0	39	2	Q29360	Q29360 sus scrofa
68	27	27.0	39	2	Q85S36	Q85S36 philoria sp
69	27	27.0	39	2	Q85S37	Q85S37 philoria pu
70	27	27.0	40	1	PHAC_MASLA	PHAC_MASLA
71	27	27.0	40	1	Y041_TREPA	Y041_TREPA
72	27	27.0	40	2	Q7RJ54	Q7RJ54 plasmodium
73	27	27.0	40	2	Q85CF5	Q85CF5 philoria pu
74	27	27.0	40	2	Q94IS0	Q94IS0 pinus radia
75	27	27.0	40	2	Q9LLL4	Q9LLL4 pyrus pyrif
76	26.5	26.5	39	2	Q9S9G2	Q9S9G2 cucumis sat
77	26.5	26.5	40	2	Q9PC59	Q9PC59 xylella fas
78	26	26.0	18	2	Q9QV11	Q9QV11 rattus sp.
79	26	26.0	25	2	Q9QVA3	Q9QVA3 rattus sp.
80	26	26.0	29	2	Q96BS8	Q96BS8 homo sapien
81	26	26.0	30	2	Q981L4	Q981L4 rhizobium l
82	26	26.0	32	2	Q9IWC2	Q9IWC2 phocoena sp
83	26	26.0	33	2	Q8WWE0	Q8WWE0 homo sapien
84	26	26.0	33	2	Q6JU23	Q6JU23 ovis aries
85	26	26.0	33	2	Q82Y96	Q82Y96 streptomyc
86	26	26.0	34	2	Q9H530	Q9H530 homo sapien
87	26	26.0	34	2	Q8EY20	Q8EY20 leptospira
88	26	26.0	35	2	Q870I2	Q870I2 saccharomyc
89	26	26.0	36	2	Q73HNO	Q73HNO wolbachia p
90	26	26.0	36	1	TXAM_METSE	TXAM_METSE
91	26	26.0	37	1	MIBP_PSESP	MIBP_PSESP
92	26	26.0	37	2	Q8ZAP7	Q8ZAP7 versinia pe
93	26	26.0	37	2	Q7UEH1	Q7UEH1 rhodopirell
94	26	26.0	38	2	Q52382	Q52382 pseudomonas
95	26	26.0	39	2	Q8TGQ3	Q8TGQ3 saccharomyc
96	26	26.0	39	2	Q88383	Q88383 simian t-ly
97	26	26.0	40	2	Q8S5P4	Q8S5P4 oryza sativ
98	26	26.0	40	2	Q7M258	Q7M258 pium sativ
99	25.5	25.5	35	2	Q9R623	Q9R623 bacillus su
100	25	25.0	18	2	Q6LD16	Q6LD16 homo sapien

ALIGNMENTS

```
RESULT 1
O39812
ID O39812 PRELIMINARY; PRT; 33 AA.
AC O39812;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE VP3 (Fragment).
OS Encephalomyocarditis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus.
OX NCBI_TaxID=12104;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97445656; PubMed=9300541; DOI=10.1016/S0034-5288(97)90197-1;
RA Koenen F., Papadopoulos O., Paschaleri-Papadopolou E., Brocchi E.,
RA De, Simone F., Carra E., Knowles N.J.;
RT "Comparison of the pathogenic, antigenic and molecular characteristics
of two encephalomyocarditis virus (EMCV) isolates from Belgium and
Greece.";
RT Res. Vet. Sci. 62:239-244(1997).
DR EMBL; Y14456; CAA74801.1; -.
DR HSSP; P12296; 2MEV.
DR GO; GO:0019028; C:viral capsid; IEA.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; Rhv; 1.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3464 MW; BD67B5C5B27AFF20 CRC64;

Query Match 36.0%; Score 36; DB 2; Length 33;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GTDMKRLRLPASP 13
DB 16 GKDFSLKMPISP 27

RESULT 2
O39814
ID O39814 PRELIMINARY; PRT; 33 AA.
AC O39814;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE VP3 (Fragment).
OS Encephalomyocarditis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus.
OX NCBI_TaxID=12104;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97445656; PubMed=9300541; DOI=10.1016/S0034-5288(97)90197-1;
RA Koenen F., Papadopoulos O., Paschaleri-Papadopolou E., Brocchi E.,
RA De, Simone F., Carra E., Knowles N.J.;
RT "Comparison of the pathogenic, antigenic and molecular characteristics
of two encephalomyocarditis virus (EMCV) isolates from Belgium and
Greece.";
RT Res. Vet. Sci. 62:239-244(1997).
DR EMBL; Y14457; CAA74803.1; -.
DR HSSP; P12296; 2MEV.
DR GO; GO:0019028; C:viral capsid; IEA.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; Rhv; 1.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3483 MW; 516FBE45B26DD591 CRC64;

Query Match 36.0%; Score 36; DB 2; Length 33;

Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GTDMKRLRLPASP 13
DB 16 GKDFSLKMPISP 27

RESULT 3
O39816
ID O39816 PRELIMINARY; PRT; 33 AA.
AC O39816;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE VP3 (Fragment).
OS Encephalomyocarditis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus.
OX NCBI_TaxID=12104;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97445656; PubMed=9300541; DOI=10.1016/S0034-5288(97)90197-1;
RA Koenen F., Papadopoulos O., Paschaleri-Papadopolou E., Brocchi E.,
RA De, Simone F., Carra E., Knowles N.J.;
RT "Comparison of the pathogenic, antigenic and molecular characteristics
of two encephalomyocarditis virus (EMCV) isolates from Belgium and
Greece.";
RT Res. Vet. Sci. 62:239-244(1997).
DR EMBL; Y14458; CAA74805.1; -.
DR HSSP; P12296; 2MEV.
DR GO; GO:0019028; C:viral capsid; IEA.
DR InterPro; IPR001676; Rhv.
DR Pfam; PF00073; Rhv; 1.
FT NON_TER 1
FT NON_TER 33
SQ SEQUENCE 33 AA; 3531 MW; 255CFBC2C20DFF3E CRC64;

Query Match 36.0%; Score 36; DB 2; Length 33;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GTDMKRLRLPASP 13
DB 16 GKDFSLKMPISP 27

RESULT 4
O39817
ID O39817 PRELIMINARY; PRT; 35 AA.
AC O39817;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=Xf0191;
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9a5c;
RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Barroga R., Alves J.E., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.P., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 19:48:47 ; Search time 46.589 Seconds
(without alignments)
140.976 Million cell updates/sec

Title: US-09-632-036f-1

Perfect score: 100

Sequence: 1 TGTDMKLRLPASPETHLDM 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 394643

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	32	14	US-10-356-824-2
2	100	100.0	32	14	US-10-406-925-2
3	100	100.0	32	14	US-10-429-519-2
4	100	100.0	32	15	US-10-356-824-2
5	100	100.0	32	15	US-10-600-152-2
6	100	100.0	32	16	US-10-909-998-2
7	80	80.0	15	15	US-10-149-138-3771
8	80	80.0	15	15	US-10-149-138-4452
9	80	80.0	15	16	US-10-149-138-3771
10	80	80.0	15	16	US-10-149-138-4452
11	79	79.0	15	15	US-10-149-138-3741
12	79	79.0	15	15	US-10-149-138-4453
13	79	79.0	15	16	US-10-149-138-3741

14	79	79.0	15	16	US-10-149-138-4453	Sequence 4453, Ap
15	64	64.0	15	15	US-10-149-138-3951	Sequence 3951, Ap
16	64	64.0	15	15	US-10-149-138-4508	Sequence 4508, Ap
17	64	64.0	15	16	US-10-149-138-3951	Sequence 3951, Ap
18	64	64.0	15	16	US-10-149-138-4508	Sequence 4508, Ap
19	59	59.0	11	15	US-10-149-138-1515	Sequence 1515, Ap
20	59	59.0	11	15	US-10-149-138-2164	Sequence 2164, Ap
21	59	59.0	11	16	US-10-149-138-1515	Sequence 1515, Ap
22	59	59.0	11	16	US-10-149-138-2164	Sequence 2164, Ap
23	58	58.0	11	15	US-10-149-138-2777	Sequence 2777, Ap
24	58	58.0	11	15	US-10-149-138-3395	Sequence 3395, Ap
25	58	58.0	11	16	US-10-149-138-2777	Sequence 2777, Ap
26	58	58.0	11	16	US-10-149-138-3395	Sequence 3395, Ap
27	57	57.0	11	15	US-10-149-138-1668	Sequence 1668, Ap
28	57	57.0	11	16	US-10-149-138-1668	Sequence 1668, Ap
29	56	56.0	11	15	US-10-149-138-3073	Sequence 3073, Ap
30	56	56.0	11	16	US-10-149-138-3073	Sequence 3073, Ap
31	55	55.0	10	15	US-10-149-138-569	Sequence 569, App
32	55	55.0	10	15	US-10-149-138-1920	Sequence 1920, Ap
33	55	55.0	10	16	US-10-149-138-569	Sequence 569, App
34	55	55.0	10	16	US-10-149-138-1920	Sequence 1920, Ap
35	55	55.0	11	15	US-10-149-138-570	Sequence 570, App
36	55	55.0	11	15	US-10-149-138-1921	Sequence 1921, Ap
37	55	55.0	11	16	US-10-149-138-570	Sequence 570, App
38	55	55.0	11	16	US-10-149-138-1921	Sequence 1921, Ap
39	53	53.0	10	15	US-10-117-937-332	Sequence 332, App
40	53	53.0	10	15	US-10-149-138-673	Sequence 673, App
41	53	53.0	10	15	US-10-149-138-1337	Sequence 1337, Ap
42	53	53.0	10	16	US-10-149-138-673	Sequence 673, App
43	53	53.0	10	16	US-10-149-138-1337	Sequence 1337, Ap
44	51	51.0	10	15	US-10-117-937-326	Sequence 326, App
45	51	51.0	10	15	US-10-149-138-349	Sequence 349, App
46	51	51.0	10	16	US-10-149-138-2700	Sequence 2700, Ap
47	51	51.0	10	16	US-10-149-138-349	Sequence 349, App
48	51	51.0	10	16	US-10-149-138-2700	Sequence 2700, Ap
49	50	50.0	10	15	US-10-149-138-417	Sequence 417, App
50	50	50.0	10	16	US-10-149-138-417	Sequence 417, App
51	49	49.0	9	15	US-10-149-138-3015	Sequence 3015, Ap
52	49	49.0	9	15	US-10-149-138-3543	Sequence 3543, Ap
53	49	49.0	9	16	US-10-149-138-3015	Sequence 3015, Ap
54	49	49.0	9	16	US-10-149-138-3543	Sequence 3543, Ap
55	48	48.0	9	15	US-10-117-937-331	Sequence 331, App
56	48	48.0	9	15	US-10-149-138-1514	Sequence 1514, Ap
57	48	48.0	9	15	US-10-149-138-1742	Sequence 1742, Ap
58	48	48.0	9	16	US-10-149-138-4185	Sequence 4185, Ap
59	48	48.0	9	16	US-10-149-138-1514	Sequence 1514, Ap
60	48	48.0	9	16	US-10-149-138-1742	Sequence 1742, Ap
61	48	48.0	9	16	US-10-149-138-4185	Sequence 4185, Ap
62	48	48.0	10	15	US-10-117-937-334	Sequence 334, App
63	48	48.0	10	15	US-10-149-138-1743	Sequence 1743, Ap
64	48	48.0	10	16	US-10-149-138-1743	Sequence 1743, Ap
65	48	48.0	11	15	US-10-149-138-852	Sequence 852, App
66	48	48.0	11	15	US-10-149-138-2486	Sequence 2486, Ap
67	48	48.0	11	15	US-10-149-138-3212	Sequence 3212, Ap
68	48	48.0	11	16	US-10-149-138-852	Sequence 852, App
69	48	48.0	11	16	US-10-149-138-2486	Sequence 2486, Ap
70	48	48.0	11	16	US-10-149-138-3212	Sequence 3212, Ap
71	47	47.0	10	15	US-10-149-138-4259	Sequence 4259, Ap
72	47	47.0	10	16	US-10-149-138-4259	Sequence 4259, Ap
73	45	45.0	9	15	US-10-149-138-3066	Sequence 3066, Ap
74	45	45.0	9	15	US-10-149-138-3975	Sequence 3975, Ap
75	45	45.0	9	16	US-10-149-138-4005	Sequence 4005, Ap
76	45	45.0	9	16	US-10-149-138-3066	Sequence 3066, Ap
77	45	45.0	9	16	US-10-149-138-3975	Sequence 3975, Ap
78	45	45.0	9	16	US-10-149-138-4005	Sequence 4005, Ap
79	45	45.0	10	15	US-10-149-138-188	Sequence 188, App
80	45	45.0	10	15	US-10-149-138-1068	Sequence 1068, Ap
81	45	45.0	10	15	US-10-149-138-1764	Sequence 1764, Ap
82	45	45.0	10	16	US-10-149-138-188	Sequence 188, App
83	45	45.0	10	16	US-10-149-138-1068	Sequence 1068, Ap
84	45	45.0	10	16	US-10-149-138-1764	Sequence 1764, Ap
85	44	44.0	8	15	US-10-117-937-330	Sequence 330, App
86	44	44.0	8	15	US-10-149-138-568	Sequence 568, App

Sequence 1571, Ap
Sequence 1919, Ap
Sequence 2305, Ap
Sequence 568, App
Sequence 1571, Ap
Sequence 1919, Ap
Sequence 2305, Ap
Sequence 333, App
Sequence 1572, Ap
Sequence 1572, Ap
Sequence 672, App
Sequence 3072, Ap
Sequence 3577, Ap
Sequence 672, App

87 44 44.0 8 15 US-10-149-138-1571
88 44 44.0 8 15 US-10-149-138-1919
89 44 44.0 8 15 US-10-149-138-2305
90 44 44.0 8 16 US-10-149-138-568
91 44 44.0 8 16 US-10-149-138-1571
92 44 44.0 8 16 US-10-149-138-1919
93 44 44.0 8 16 US-10-149-138-2305
94 44 44.0 9 15 US-10-117-937-333
95 44 44.0 9 15 US-10-149-138-1572
96 44 44.0 9 16 US-10-149-138-1572
97 41 41.0 8 15 US-10-149-138-672
98 41 41.0 8 15 US-10-149-138-3072
99 41 41.0 8 15 US-10-149-138-3577
100 41 41.0 8 16 US-10-149-138-672

ALIGNMENTS

RESULT 1
US-10-356-824-2
; Sequence 2, Application US/10356824
; Publication NO. US20030147894A1
; GENERAL INFORMATION:
; APPLICANT: Shak, Steve
; APPLICANT: Paton, Virginia
; TITLE OF INVENTION: TREATMENT WITH ANTI-ErbB2 ANTIBODIES
; FILE REFERENCE: P1256R1
; CURRENT APPLICATION NUMBER: US/10/356,824
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/208,649
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/069,346
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-356-824-2

Query Match 100.0%; Score 100; DB 14; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTDMKRLPASPETHLDM 19
Db 6 TGTDMKRLPASPETHLDM 24

RESULT 2
US-10-406-925-2
; Sequence 2, Application US/10406925
; Publication NO. US20030170234A1
; GENERAL INFORMATION:
; APPLICANT: Shak, Steve
; APPLICANT: Paton, Virginia
; TITLE OF INVENTION: TREATMENT WITH ANTI-ErbB2 ANTIBODIES
; FILE REFERENCE: P1256R3
; CURRENT APPLICATION NUMBER: US/10/406,925
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: US/09/209,023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/069,346
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-406-925-2

Query Match 100.0%; Score 100; DB 14; Length 32;

Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTDMKRLPASPETHLDM 19
Db 6 TGTDMKRLPASPETHLDM 24

RESULT 3
US-10-429-519-2
; Sequence 2, Application US/10429519
; Publication NO. US20030170235A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Robert
; TITLE OF INVENTION: TREATMENT WITH ANTI-ErbB2 ANTIBODIES
; FILE REFERENCE: P1757R1
; CURRENT APPLICATION NUMBER: US/10/429,519
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/568,322
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: US 60/134,085
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-429-519-2

Query Match 100.0%; Score 100; DB 14; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTDMKRLPASPETHLDM 19
Db 6 TGTDMKRLPASPETHLDM 24

RESULT 4
US-10-356-824-2
; Sequence 2, Application US/10356824
; Publication NO. US20040037823A9
; GENERAL INFORMATION:
; APPLICANT: Shak, Steve
; APPLICANT: Paton, Virginia
; TITLE OF INVENTION: TREATMENT WITH ANTI-ErbB2 ANTIBODIES
; FILE REFERENCE: P1256R1
; CURRENT APPLICATION NUMBER: US/10/356,824
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: US/09/208,649
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/069,346
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 9
; SEQ ID NO 2
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-356-824-2

Query Match 100.0%; Score 100; DB 15; Length 32;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTDMKRLPASPETHLDM 19
Db 6 TGTDMKRLPASPETHLDM 24

RESULT 5
US-10-600-152-2
; Sequence 2, Application US/10600152
; Publication NO. US20040037824A1

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 19:26:13 ; Search time 10.9315 Seconds
(without alignments)
167.234 Million cell updates/sec

Title: US-09-632-036f-1

Perfect score: 100

Sequence: 1 TGTDMKRLPASPETHLDM 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 9207

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 79:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	37.0	23	PS0273	hypothetical prote
2	35	35.0	34	B41397	hypothetical prote
3	34	34.0	35	D82835	hypothetical prote
4	31	31.0	31	I39679	exeE protein - Aer
5	30	30.0	20	JN0252	water-soluble 35K
6	30	30.0	31	T36022	small hypothetical
7	29.5	29.5	28	I54183	cell adhesion regu.
8	29	29.0	33	S26859	chitinase (EC 3.2.
9	28	28.0	20	I54189	galactose-1-phosph
10	28	28.0	22	S47195	T-cell receptor J-
11	28	28.0	23	S47208	T-cell receptor J-
12	28	28.0	23	S47203	T-cell receptor J-
13	28	28.0	25	S47211	T-cell receptor J-
14	28	28.0	33	G81837	hypothetical prote
15	27	27.0	20	A38566	GTP-binding protei
16	27	27.0	26	B28378	replication contro
17	27	27.0	28	S66436	allophycocyanin al
18	27	27.0	31	G82071	hypothetical prote
19	27	27.0	32	C46107	polymavirus enhan
20	27	27.0	40	B27398	allophycocyanin al
21	27	27.0	40	H71373	hypothetical prote
22	26.5	26.5	40	G82620	hypothetical prote
23	26	26.0	19	S69166	hypothetical prote
24	26	26.0	22	A39269	ferredoxin b - Jap
25	26	26.0	35	S43485	methylntransferase
26	26	26.0	36	A27222	metridin (EC 3.4.2
27	26	26.0	37	AH0455	hypothetical prote
28	26	26.0	38	D55543	Tnp protein - Pseu
29	26	26.0	39	B31940	microtubule-associ

30	26	26.0	40	2	S09338	ferredoxin [2Fe-2S
31	26	26.0	40	2	S71301	ICL5 protein - Par
32	25	25.0	18	2	G42753	interferon alpha (
33	25	25.0	26	2	H42753	interferon alpha (
34	25	25.0	28	2	I46921	gene Bota protein
35	25	25.0	32	2	S39628	probable urease (E
36	25	25.0	32	2	G84161	hypothetical prote
37	25	25.0	33	2	I38197	hypothetical prote
38	25	25.0	35	2	S04182	hypothetical prote
39	25	25.0	36	2	F42753	interferon alpha (
40	25	25.0	36	2	B31485	Ig heavy chain V r
41	25	25.0	37	2	B24553	cuticle protein SC
42	25	25.0	39	2	C24553	cuticle protein SC
43	25	25.0	40	2	S33403	Ig heavy chain V r
44	24	24.0	25	2	PC3489	sulfite reductase
45	24	24.0	26	2	B46451	RNA-directed RNA p
46	24	24.0	26	2	D32248	Ig kappa chain V r
47	24	24.0	26	2	JT0965	cytochrome-c oxida
48	24	24.0	27	2	S51176	aspartate transcar
49	24	24.0	30	2	A44912	cysteine proteinas
50	24	24.0	31	2	S32610	antiviral protein
51	24	24.0	32	2	G31485	Ig heavy chain V r
52	24	24.0	32	2	B38839	histone H1 - sea u
53	24	24.0	33	2	C82419	hypothetical prote
54	24	24.0	34	2	A54001	superoxide dismuta
55	24	24.0	35	1	LNFGC1	pulmonary surfacta
56	24	24.0	35	2	S11614	ribosomal protein
57	24	24.0	39	2	A31763	neural cell adhesi
58	24	24.0	39	2	A60716	somatotropin intro
59	24	24.0	39	2	C83904	hypothetical prote
60	24	24.0	40	2	S71295	deoxyguanosine kin
61	24	24.0	40	2	G69471	hypothetical prote
62	24	24.0	40	2	T07583	hypothetical prote
63	23.5	23.5	30	2	E45095	photosystem I ligh
64	23.5	23.5	30	2	C71309	hypothetical prote
65	23.5	23.5	37	2	S29113	dipteridin homolog
66	23.5	23.5	38	2	T34678	hypothetical prote
67	23	23.0	14	2	H64008	hypothetical prote
68	23	23.0	15	2	PQ0545	ermK leader peptid
69	23	23.0	15	2	S54271	capsid protein VP1
70	23	23.0	16	2	S54271	GATA-2 protein - A
71	23	23.0	17	2	S17274	ribosomal protein
72	23	23.0	19	2	B61409	genome polyprotein
73	23	23.0	20	2	S56756	link protein - rat
74	23	23.0	20	2	S38288	50K allergen - per
75	23	23.0	23	2	S43289	cytochrome-c oxida
76	23	23.0	24	2	I54329	gene NF1 protein -
77	23	23.0	25	2	A61148	cyanidase (EC 3.5.
78	23	23.0	26	2	A32806	transposase - Esch
79	23	23.0	27	2	A49046	T-cell receptor al
80	23	23.0	28	2	B60583	glycoprotein hormo
81	23	23.0	28	2	S56121	type I DNA methylt
82	23	23.0	29	2	C61233	conceptus protein
83	23	23.0	30	2	A84412	hypothetical prote
84	23	23.0	31	2	S03297	Ig alpha chain C r
85	23	23.0	31	2	S78771	camelysin COMP - B
86	23	23.0	31	2	A59357	neurotoxin Bmk KTX
87	23	23.0	31	2	A59321	protein BmSKTx2 [i
88	23	23.0	32	2	S03352	malate dehydrogena
89	23	23.0	35	2	F81179	hypothetical prote
90	23	23.0	35	2	C69977	hypothetical prote
91	23	23.0	36	1	YFG	peptide YY - pig
92	23	23.0	36	2	A60416	peptide YY - dog
93	23	23.0	36	2	A37172	collagen alpha 1(X
94	23	23.0	37	1	A42040	kallitoxin 1 [vali
95	23	23.0	37	2	A26781	peptide YG - Ameri
96	23	23.0	37	2	I48405	histone H2a - mous
97	23	23.0	37	2	B39030	androgen-binding p
98	23	23.0	37	2	A86127	hypothetical prote
99	23	23.0	39	2	B70162	ribosomal protein
100	23	23.0	40	2	I49405	beta-galactoside-b

ALIGNMENTS

RESULT 1
PS0273
hypothetical protein 3 (repB 3' region) - Thiobacillus ferrooxidans plasmid pTF-FC2 (fr
C:Species: Thiobacillus ferrooxidans
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 18-Jun-1993
R:Accession: PS0273
R:Dorington, R.A.; Bardien, S.; Rawlings, D.E.
Gene 108, 7-14, 1991

A:Title: The broad-host-range plasmid pTF-FC2 requires a primase-like protein for autono
A:Reference number: JS0637; MUID:92104508; PMID:1761233
A:Accession: PS0273
A:Molecule type: DNA
A:Residues: 1-23 <DOR>
A:Cross-references: GB:M64901
C:Genetics:
A:Genome: plasmid

Query Match 37.0%; Score 37; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 MKRLPASPETHLD 18
: : ||||| :
Db 2 LAIRLPAEVETRL 15

RESULT 2
B41397
hypothetical protein 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 03-May-1994
R:Accession: B41397
R:Miller, D.A.; Lee, A.; Matsui, Y.; Chen, E.Y.; Moses, H.L.; Derynck, R.
Mol. Endocrinol. 3, 1926-1934, 1989
A:Title: Complementary DNA cloning of the murine transforming growth factor-beta3 (TGFB
t tissues.

A:Reference number: A41397; MUID:90190650; PMID:2628730
A:Accession: B41397
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-34 <MIL>
A:Cross-references: GB:M32745

Query Match 35.0%; Score 35; DB 2; Length 34;
Best Local Similarity 75.0%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 LPASPETH 16
: : ||||| :
Db 24 LPGSPEAH 31

RESULT 3
D82835
hypothetical protein XF0191 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
R:Accession: D82835
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below

A:Accession: D82835
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-35 <SIM>
A:Cross-references: UNIPROT:O9PGV7; GB:AE003873; GB:AE003849; NID:g9104992; PIDN:AAF8300
A:Experimental source: strain 9a5c
R:Simpsom, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.B.; Marques, M.V.; Martins, E.
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0191

Query Match 34.0%; Score 34; DB 2; Length 35;
Best Local Similarity 28.6%; Pred. No. 67;
Matches 4; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 3 TDMKRLPASPETH 16
: : ||||| :
Db 11 SNIKIPNPSLH 24

RESULT 4
I39679
exeE protein - Aeromonas salmonicida (fragment)
C:Species: Aeromonas salmonicida
C:Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004
R:Accession: I39679
R:Karlyshev, A.V.; MacIntyre, S.
Gene 158, 77-82, 1995
A:Title: Cloning and study of the genetic organization of the exe gene cluster of Aeromori
A:Reference number: I39675; MUID:95309729; PMID:7789814
A:Accession: I39679
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-31 <RES>
A:Cross-references: UNIPROT:Q44263; EMBL:X80505; NID:gs16349; PIDN:CAA56669.1; PID:g10174

C:Superfamily: secretion protein xcpR

Query Match 31.0%; Score 31; DB 2; Length 31;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 GTDMKRLPASP 13
: : ||||| :
Db 8 GTDLPALPELP 19

RESULT 5
JN0252
water-soluble 35K protein - soybean (fragment)
C:Species: Glycine max (soybean)
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 09-Jul-2004
R:Accession: JN0252
R:Akhtaruzzaman, M.; Kimura, Y.; Takagi, S.
Biosci. Biotechnol. Biochem. 56, 873-878, 1992
A:Title: Glycinin A4A5 subunit digesting protease in soybean seeds.
A:Reference number: JN0252
A:Accession: JN0252
A:Molecule type: protein
A:Residues: 1-20 <AKH>
A:Cross-references: UNIPROT:Q7M212
A:Experimental source: seed
C:Superfamily: glycinin

Query Match 30.0%; Score 30; DB 2; Length 20;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 19:03:10 ; Search time 61.1644 Seconds
(without alignments)
120.143 Million cell updates/sec

Title: US-09-632-036F-1

Perfect score: 100

Sequence: 1 TGTDMKRLRPASPETHLDM 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 888314

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp19808:*

2: Geneseqp19908:*

3: Geneseqp20008:*

4: Geneseqp20018:*

5: Geneseqp20028:*

6: Geneseqp20038:*

7: Geneseqp20038s:*

8: Geneseqp20048s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	19	4	AAB586622 HER-2 B C
2	100	100.0	32	2	AAM59346 Human Erb
3	100	100.0	32	2	AAY14575 Human erb
4	100	100.0	32	4	AAB62075 ErbB2 7C2
5	100	100.0	32	4	AAB48764 Human Erb
6	80	80.0	15	4	AAG89010 Human Erb
7	80	80.0	15	4	AAG88392 HER2/NEU
8	79	79.0	15	3	AAY98841 HLA class
9	79	79.0	15	4	AAG89011 HER2/neu
10	79	79.0	15	4	AAG88332 HER2/NEU
11	78	78.0	25	4	AAG66872 Human Erb
12	78	78.0	15	3	AAY98840 HLA class
13	72	72.0	15	4	AAG66881 Human Erb
14	72	72.0	20	4	AAG66880 Human Erb
15	72	72.0	25	4	AAG66871 Human Erb
16	64	64.0	15	3	AAY98932 HLA class
17	64	64.0	15	4	AAG9066 HER2/neu
18	64	64.0	15	4	AAG88752 HER2/NEU
19	53	53.0	10	2	AAR61633 Peptide f
20	53	53.0	10	6	ABP74448 Human Her
21	53	53.0	10	7	ADC09307 Epitope w
22	53	53.0	10	8	ADE97864 Immunogen
23	51	51.0	10	6	ABP74442 Human Her
24	51	51.0	10	7	ADC09301 Epitope w
25	48	48.0	9	4	AAG88751 HER2/NEU

26	48	48.0	9	4	AAB586646	Aab586645 HER-2 CTL
27	48	48.0	9	6	ABP74447	Abp74447 Human Her
28	48	48.0	9	7	ADC09306	Adc09306 Epitope w
29	48	48.0	10	6	ABP74450	Abp74450 Human Her
30	48	48.0	10	7	ADC09309	Adc09309 Epitope w
31	47	47.0	10	4	AAG88919	Aag88919 HER2/neu
32	45	45.0	9	4	AAG88391	Aag88391 HER2/NEU
33	45	45.0	9	4	AAG88331	Aag88331 HER2/NEU
34	45	45.0	15	4	AAG66884	Aag66884 Human Erb
35	45	45.0	20	4	AAG66882	Aag66882 Human Erb
36	45	45.0	25	4	AAG66885	Aag66885 Human Erb
37	44	44.0	8	6	ABP74446	Abp74446 Human Erb
38	44	44.0	8	7	ADC09305	Adc09305 Epitope w
39	44	44.0	9	6	ABP74449	Abp74449 Human Her
40	44	44.0	9	7	ADC09308	Adc09308 Epitope w
41	40	40.0	7	8	ABP74441	Abp74441 Human HLA
42	40	40.0	8	6	ABP74441	Abp74441 Human Her
43	40	40.0	8	7	ADC09300	Adc09300 Epitope w
44	40	40.0	8	8	ADN64551	Adn64551 HLA bindi
45	40	40.0	8	8	ADP80286	Adp80286 Human HLA
46	40	40.0	11	8	ADN64641	Adn64641 HLA bindi
47	40	40.0	11	8	ADP80287	Adp80287 Human HLA
48	40	40.0	18	4	ABB41449	Abb41449 Peptide #
49	40	40.0	18	4	AAM62323	Aam62323 Human bra
50	38	38.0	33	8	ADS84891	Ads84891 Beta-site
51	37	37.0	14	4	AAM53151	Aam53151 Human con
52	36	36.0	31	4	AAU31251	Aau31251 Novel hum
53	35	35.0	36	8	ADJ25789	Adj25789 MDM2 bind
54	34	34.0	21	4	AAM20779	Aam20779 Peptide #
55	34	34.0	21	4	ABB42333	Abb42333 Peptide #
56	34	34.0	21	4	AAM16139	Aam16139 Peptide #
57	34	34.0	21	4	ABB25824	Abb25824 Protein #
58	34	34.0	21	4	AAM76032	Aam76032 Human bon
59	34	34.0	21	4	AAM63219	Aam63219 Human bra
60	34	34.0	21	4	ABG57758	Abg57758 Human liv
61	34	34.0	21	5	ABG45430	Abg45430 Human pep
62	34	34.0	27	5	AAU89413	Aau89413 Insulin/i
63	34	34.0	33	6	ABU08630	Abu08630 LS147-enc
64	34	34.0	37	8	ABO53961	Abo53961 Human gen
65	33	33.0	10	6	ABP74445	Abp74445 Human Her
66	33	33.0	10	7	ADC09304	Adc09304 Epitope w
67	33	33.0	11	4	AAG88920	Aag88920 HER2/neu
68	33	33.0	15	6	ADA00906	Ada00906 Mouse GL
69	33	33.0	24	7	ADC27958	Adc27958 Synthetic
70	33	33.0	29	8	ABO57829	Abo57829 Human gen
71	33	33.0	30	5	ABG62068	Abg62068 Human pro
72	33	33.0	37	8	AAO90919	Aao90919 Human pol
73	33	33.0	37	8	ADJ45726	Adj45726 Signal pe
74	33	33.0	38	2	AAR58380	Aar58380 TSAR bind
75	32	32.0	9	2	AAY37971	Aay37971 Human cer
76	32	32.0	14	2	AAY45541	Aay45541 Immunogen
77	32	32.0	14	4	AAM53150	Aam53150 Human con
78	32	32.0	15	2	AAR97899	Aar97899 Japan ced
79	32	32.0	16	5	ABB06398	Abb06398 Intracell
80	32	32.0	24	3	AAY58888	Aay58888 Polyketid
81	32	32.0	24	7	ADC27959	Adc27959 Synthetic
82	32	32.0	26	2	AAI12216	Aai12216 Human 5'
83	32	32.0	27	8	ADP81213	Adp81213 Protein o
84	32	32.0	32	3	ADQ95637	Adq95637 Aggrecon
85	32	32.0	34	3	AAB30463	Aab30463 Amino aci
86	32	32.0	35	4	AAM15419	Aam15419 Peptide #
87	32	32.0	35	4	ABB34431	Abb34431 Peptide #
88	32	32.0	35	4	AAM27909	Aam27909 Peptide #
89	32	32.0	35	4	ABB19841	Abb19841 Protein #
90	32	32.0	35	4	AAM67613	Aam67613 Human bon
91	32	32.0	35	4	AAM55217	Aam55217 Human bra
92	32	32.0	35	4	AAM03181	Aam03181 Peptide #
93	32	32.0	39	5	ABB06391	Abb06391 Human int
94	32	32.0	39	5	ABB06392	Abb06392 Carp intr
95	32	32.0	39	5	ABB06406	Abb06406 Intracell
96	32	32.0	40	3	AAG33986	Aag33986 Arabidops
97	32	32.0	40	5	ABB06397	Abb06397 Intracell
98	31	31.0	12	2	AAR23579	Aar23579 HTLV-II d

99 31 31.0 13 8 ADM12693 Ii key/WH
100 31 31.0 13 8 ADO38929 Anthrax t

ALIGNMENTS

```
RESULT 1
AAB68622
ID AAB68622 standard; peptide; 19 AA.
XX
AC AAB68622;
XX
DT 30-APR-2001 (first entry)
XX
DE HER-2 B cell epitope #1.
XX
KW Cytostatic; immune response; HER-2; human; epitope; cancer; breast;
KW ovarian; lung; prostate; colon.
XX
OS Homo sapiens.
XX
PN WO200108636-A2.
XX
PD 08-FEB-2001.
XX
PF 03-AUG-2000; 2000WO-US021222.
XX
PR 03-AUG-1999; 99US-0146869P.
XX
PA (OHIS ) UNIV OHIO STATE.
XX
PI Kaumaya PT, Stevens VC, Triozzi PL;
XX
DR WPI; 2001-182849/18.
XX
PT Compositions comprising polypeptides and polynucleotides for stimulating
PT the immune system and for treating malignancies associated with
PT overexpression of the HER-2 protein.
XX
PS Claim 1; Page 37; 51pp; English.
XX
CC The present invention relates to compositions for stimulating the immune
CC system and for treating malignancies associated with overexpression of
CC the HER-2 protein. The compositions comprise immunogenic groups of the
CC HER-2 proteins. The present sequence is one such peptide used in the
CC compositions of the present invention. The compositions can be used for
CC treating cancer, e.g. breast, ovarian, lung, prostate and colon cancers
XX
SQ Sequence 19 AA;
Query Match 100.0%; Score 100; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTDMKRLPASPETHLDM 19
Db 1 TGTDMKRLPASPETHLDM 19
RESULT 2
AAW59346
ID AAW59346 standard; protein; 32 AA.
XX
AC AAW59346;
XX
DT 08-SEP-1998 (first entry)
XX
DE Human ErbB2 extracellular domain protein fragment (aa 22-53).
XX
KW ErbB2; epitope; monoclonal antibody; identification; apoptosis;
KW detection; treatment; disorder; tumour; benign; malignant; leukaemia;
KW lymphoid malignancy; inflammation; tumour; angiogenic; immunological.
```

```
XX Homo sapiens.
OS WO9817797-A1.
XX
PN 30-APR-1998.
XX
PF 09-OCT-1997; 97WO-US018385.
XX
PR 18-OCT-1996; 96US-00731794.
XX
PA (GETH ) GENENTECH INC.
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
PI Fendly BM, Phillips GD, Scheuermann RH, Uhr JW;
XX WPI; 1998-261493/23.
XX
DR New anti-ErbB2 antibodies - which induce apoptosis in cells which
XX overexpress ErbB2, used to treat e.g. tumours or inflammatory, angiogenic
PT or immunologic disorders.
PT
XX Disclosure; Page 45; 72pp; English.
XX
CC This sequence represents a fragment of the human ErbB2 extracellular
CC domain corresponding to amino acids 22-53 which is capable of binding the
CC monoclonal antibodies 7C3 and 7F3. This region is used in a method to
CC identify antibodies which bind to ErbB2 and induce apoptosis of a cell
CC which overexpresses ErbB2. The antibodies can also be used to detect
CC ErbB2 on a cell. The antibodies can be used for treating disorders such
CC as benign or malignant tumours, (e.g. renal, liver, kidney, bladder,
CC breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
CC thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head
CC and neck tumours), leukemia's and lymphoid malignancies, other disorders
CC such as neuronal, glial, astrocytic, hypothalamic and other glandular,
CC macrophagal, epithelial, stromal and blastocoeic disorders, and
CC inflammatory, angiogenic and immunologic disorders
XX
SQ Sequence 32 AA;
Query Match 100.0%; Score 100; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTDMKRLPASPETHLDM 19
Db 6 TGTDMKRLPASPETHLDM 24
RESULT 3
AAV14575
ID AAV14575 standard; peptide; 32 AA.
XX
AC AAV14575;
XX
DT 14-SEP-1999 (first entry)
XX
DE Human erbB2 receptor-derived epitope 7C2/7F3.
XX
KW Domain; human; erbB2; receptor; antibody; overexpression; anthracycline;
KW chemotherapeutic agent; tumour; cancer; breast; squamous cell; cervical;
KW non-small-cell lung; gastrointestinal; pancreatic; glioblastoma; ovarian;
KW liver; bladder; hepatoma; colon; colorectal; endometrial carcinoma;
KW salivary gland; kidney; prostate; vulva; thyroid; hepatic carcinoma;
KW head-and-neck; epitope.
XX
OS Homo sapiens.
XX
PN WO9931140-A1.
XX
PD 24-JUN-1999.
XX
PF 10-DEC-1998; 98WO-US026266.
```



```
/clone.lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```
Alignment Scores:
Pred. No.:      8,49e+03      Length:      40
Score:          36.00         Matches:      8
Percent Similarity: 72.73%      Conservative: 0
Best Local Similarity: 72.73%      Mismatches:  3
Query Match:    36.00%         Indels:      0
DB:             2             Gaps:        0
```

US-09-632-036F-1 (1-19) x BE733350 (1-40)

```
Qy      3 ThrAspMetLysLeuArgLeuProAlaSerPro 13
      |||||
Db      6 ACACAGCAGGAACACTACGCGTACGTGGAGGCCG 38
      |||||
```

RESULT 2

```
AL943196/c
LOCUS      AL943196          39 bp      DNA      linear      GSS      01-APR-2004
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-272D12-015091,
            genomic survey sequence.
ACCESSION  AL943196
VERSION    AL943196.1      GI:24399794
KEYWORDS   GSS.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicaceae; Arabidopsi.
```

```
1
Li,Y., Rosso,M.G., Strizhov,N., Viehoveer,P. and Weishaar,B.
GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
the identification of T-DNA insertion mutants in Arabidopsis
thaliana
Bioinformatics 19 (11), 1441-1442 (2003)
22755829
12874060
2
Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
Weishaar,B.
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
flanking sequence tag-based reverse genetics
Plant Mol. Biol. 53 (1-2), 247-259 (2003)
23117147
14756321
3
Strizhov,N., Li,Y., Rosso,M.G., Viehoveer,P., Dekker,K.A. and
Weishaar,B.
High-throughput generation of sequence indexes from T-DNA
mutagenized Arabidopsis thaliana lines
BioTechniques 35 (6), 1164-1168 (2003)
14682050
4 (bases 1 to 39)
Rosso,M.G., Li,Y., Strizhov,N. and Weishaar,B.
Direct Submission
```

```
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
flanking sequence tag-based reverse genetics
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Strizhov,N., Li,Y., Rosso,M.G., Viehoveer,P., Dekker,K.A. and
Weishaar,B.
```

```
TITLE
High-throughput generation of sequence indexes from T-DNA
mutagenized Arabidopsis thaliana lines
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
Rosso,M.G., Li,Y., Strizhov,N. and Weishaar,B.
Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer
Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence has been recovered from the left border of the T-DNA.
It indicates an insertion within the locus defined by BAC clone
MDN11. Details on the protocols used for generation of the sequence
are described in References 1-3. The sequences are generated at the
MPI for Plant Breeding Research in the context of the GABI-Kat
project. GABI-Kat is part of the German Plant Genomics program
designated 'GABI'. Information on line availability can be found
```

```
at: http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES
source
1..39
Location/Qualifiers
```

```
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone.lib="Arabidopsis thaliana T-DNA insertion lines"
/ecotype="Col-0"
/note="PCR was performed on DNA from Arabidopsis thaliana
plants (Ti) which were transformed with the T-DNA from
vector pAC161 (GenBank accession number: AJ537514). The
lines contain one or more T-DNA insertions. The DNA
fragment(s) resulting from the PCR were directly sequenced
to determine the genomic sequence flanking the insertion.
T-DNA derived sequences were removed."
```

ORIGIN

```
Alignment Scores:
Pred. No.:      1,2e+04      Length:      39
Score:          35.00         Matches:      6
Percent Similarity: 77.78%      Conservative: 1
Best Local Similarity: 66.67%      Mismatches:  2
Query Match:    35.00%         Indels:      0
DB:             9             Gaps:        0
```

US-09-632-036F-1 (1-19) x AL943196 (1-39)

```
Qy      8 ArgLeuProAlaSerProGluThrHis 16
      |||||
Db      35 AGACTTCCCGCTAACCAACGAGCAT 9
      |||||
```

RESULT 3

```
AI769631/c
LOCUS      AI769631          28 bp      mRNA      linear      EST 21-DEC-1999
DEFINITION wj25a04.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE:2403822 3',
            similar to TR:Q92466 Q92466 DBB P48.1, mRNA sequence.
ACCESSION  AI769631
VERSION    AI769631.1      GI:5236140
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
```

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 28)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lemmon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
```

```
Trace considered overall poor quality
Insert Length: 1994 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
```

```
1..28
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone.lib="IMAGE:2403822"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
```

```
FEATURES
source
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2005, 10:07:17 ; Search time 314.495 Seconds
(without alignments)
375.024 Million cell updates/sec

Title: US-09-632-036f-1

Perfect score: 100

Sequence: 1 TGTDMKRLPASPETHLDM 19

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 5878042

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO_spool_h/US09632036/runat_17062005_132235_8608/app_query.fasta_1.1791
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=40 -USER=US09632036 @CGN 1 1 1295 @runat_17062005_132235_8608 -NCFU=6
-ICV=3 -NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	35	35.0	19	16	US-10-251-117-17	Sequence 17, Appl
C 2	35	35.0	19	16	US-10-251-117-266	Sequence 266, App
C 3	35	35.0	25	21	US-10-719-900-80309	Sequence 80309, A
4	35	35.0	25	21	US-10-719-900-212555	Sequence 212555,
5	35	35.0	25	21	US-10-719-900-249478	Sequence 249478,
6	35	35.0	25	21	US-10-719-900-346374	Sequence 346374,
C 7	34	34.0	25	21	US-10-719-900-305176	Sequence 305176,
8	34	34.0	39	14	US-10-057-136-15	Sequence 15, Appl
C 9	34	34.0	39	16	US-10-265-811-9	Sequence 9, Appl
10	33	33.0	25	15	US-10-098-263B-111514	Sequence 111514,
C 11	33	33.0	25	15	US-10-098-263B-130752	Sequence 130752,
12	33	33.0	25	21	US-10-719-900-196604	Sequence 196604,
13	33	33.0	25	21	US-10-719-900-244731	Sequence 244731,
C 14	33	33.0	25	21	US-10-719-900-395120	Sequence 395120,
15	33	33.0	25	21	US-10-719-900-812465	Sequence 812465,
C 16	33	33.0	25	21	US-10-719-900-943313	Sequence 943313,
C 17	33	33.0	25	21	US-10-719-900-943314	Sequence 943314,
C 18	33	33.0	25	21	US-10-719-900-966842	Sequence 966842,
C 19	33	33.0	25	21	US-10-956-157-178532	Sequence 178532,
C 20	33	33.0	30	15	US-10-134-643-21	Sequence 21, Appl
C 21	32	32.0	19	16	US-10-251-117-15	Sequence 15, Appl
C 22	32	32.0	19	16	US-10-251-117-264	Sequence 264, App
C 23	32	32.0	25	21	US-10-719-900-44912	Sequence 44912, A
C 24	32	32.0	25	21	US-10-719-900-218229	Sequence 218229,
C 25	32	32.0	25	21	US-10-719-900-218230	Sequence 218230,
C 26	32	32.0	25	21	US-10-719-900-370643	Sequence 370643,
C 27	32	32.0	25	21	US-10-719-900-542722	Sequence 542722,
28	32	32.0	25	21	US-10-719-900-615040	Sequence 615040,
C 29	32	32.0	25	21	US-10-809-189-18778	Sequence 18778, A
C 30	32	32.0	27	14	US-10-183-728-4	Sequence 4, Appl
C 31	32	32.0	35	15	US-10-203-821-10	Sequence 10, Appl
C 32	32	32.0	37	9	US-09-944-243-4	Sequence 4, Appl
C 33	32	32.0	37	17	US-10-444-031-4	Sequence 4, Appl
C 34	32	32.0	38	9	US-09-881-823-24	Sequence 24, Appl
C 35	32	32.0	38	9	US-10-461-790-74	Sequence 74, Appl
C 36	31	31.0	24	14	US-10-024-648-2	Sequence 23129, A
C 37	31	31.0	25	15	US-10-098-263B-23129	Sequence 67215, A
C 38	31	31.0	25	21	US-10-719-900-67215	Sequence 99561, A
C 39	31	31.0	25	21	US-10-719-900-99561	Sequence 196313,
C 40	31	31.0	25	21	US-10-719-900-196313	Sequence 196314,
C 41	31	31.0	25	21	US-10-719-900-196314	Sequence 212556,
C 42	31	31.0	25	21	US-10-719-900-212556	Sequence 279126,
C 43	31	31.0	25	21	US-10-719-900-279126	Sequence 299529,
C 44	31	31.0	25	21	US-10-719-900-299529	Sequence 343564,
45	31	31.0	25	21	US-10-719-900-343564	

ALIGNMENTS

RESULT 1

US-10-251-117-17
; Sequence 17, Application US/10251117
; Publication No. US20030170891A1
; GENERAL INFORMATION:
; APPLICANT: McSwiggen, James
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; TITLE OF INVENTION: RNA Interference Mediated Inhibition of Epidermal Growth Factor R
; FILE REFERENCE: 900/042 (MEHB02-468-A)
; CURRENT APPLICATION NUMBER: US/10/251,117
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US 60/393,924
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 10/163,552
; PRIOR FILING DATE: 2002-06-06

; PRIOR APPLICATION NUMBER: US 60/358,580
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 09/916,466
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/296,249
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 1213
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Target sequence/siNA sense
US-10-251-117-17

Alignment Scores:
Pred. No.: 394 Length: 19
Score: 35.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.00% Indels: 0
DB: 16 Gaps: 0

US-09-632-036F-1 (1-19) x US-10-251-117-17 (1-19)
Qy 13 ProGluThrHisLeuAsp 18
Db 1 CCCGAGACCCACCGGAC 18

RESULT 2

US-10-251-117-266/c

; Sequence 266, Application US/10251117
; Publication No. US20030170891A1
; GENERAL INFORMATION:

; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; TITLE OF INVENTION: RNA Interference Mediated Inhibition of Epidermal Growth Factor R

; FILE REFERENCE: 900/042 (MBH02-468-A)
; CURRENT APPLICATION NUMBER: US/10/251,117
; CURRENT FILING DATE: 2003-02-24

; PRIOR APPLICATION NUMBER: US 60/393,924
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 10/163,552
; PRIOR FILING DATE: 2002-06-06

; PRIOR APPLICATION NUMBER: US 60/358,580
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 09/916,466
; PRIOR FILING DATE: 2001-07-25

; PRIOR APPLICATION NUMBER: US 60/296,249
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 1213
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 266
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: siNA antisense region
US-10-251-117-266

Alignment Scores:
Pred. No.: 394 Length: 19
Score: 35.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.00% Indels: 0
DB: 16 Gaps: 0

US-09-632-036F-1 (1-19) x US-10-251-117-266 (1-19)
Qy 13 ProGluThrHisLeuAsp 18

Db 19 CCCGAGACCCACCTGGAC 2

RESULT 3

US-10-719-900-80309/c

; Sequence 80309, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:

; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse

; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20

; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 80309
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus

US-10-719-900-80309
Alignment Scores:
Pred. No.: 525 Length: 25
Score: 35.00 Matches: 5
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 0
Query Match: 35.00% Indels: 0
DB: 21 Gaps: 0

US-09-632-036F-1 (1-19) x US-10-719-900-80309 (1-25)
Qy 9 LeuProAlaSerProGluThrHis 16
Db 25 TTGCCGAGCAATCCACAACCTCAT 2

RESULT 4

US-10-719-900-212555

; Sequence 212555, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:

; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse

; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20

; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 212555
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus

US-10-719-900-212555
Alignment Scores:
Pred. No.: 525 Length: 25
Score: 35.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 35.00% Indels: 0
DB: 21 Gaps: 0

US-09-632-036F-1 (1-19) x US-10-719-900-212555 (1-25)
Qy 12 SerProGluThrHisLeuAsp 18
Db 3 TCTCCGAGACCCCAATTGGAC 23

RESULT 5

US-10-719-900-249478

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2005, 09:06:41 ; Search time 81.5191 Seconds
(without alignments)

381.374 Million cell updates/sec

Title: US-09-632-036f-1

Perfect score: 100

Sequence: 1 TGTDMKRLPASPETHLDM 19

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Fgapop 10.0 , Fgapext 0.5

Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 1132682

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh

-Q=/cgn2_1/USPTO_spool_1/US09632036/runat_17062005_132233_8571/app_query.fasta_1.1791

-DB=issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi

-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=40

-USER=US09632036 @CGN 1.1 258 @runat_17062005_132233_8571 -NCPU=6 -ICPU=3

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	33	33.0	32	3	US-09-246-277A-3
C 2	32.5	32.5	36	5	PCT-US93-07672-5
C 3	32	32.0	18	4	US-09-663-834A-10
C 4	32	32.0	25	4	US-09-396-196G-18778
C 5	32	32.0	33	1	US-08-435-350-5
C 6	32	32.0	37	3	US-09-366-920-4
C 7	32	32.0	37	4	US-09-944-243-4
C 8	31	31.0	25	4	US-09-396-196G-16697
C 9	31	31.0	25	4	US-09-396-196G-61016
C 10	31	31.0	25	4	US-09-396-196G-109354
C 11	31	31.0	38	1	US-08-360-841-12
C 12	31	31.0	38	3	US-09-308-1608-8

13	31.0	38	3	US-09-456-399-8	Sequence 8, Appl
C 14	30.0	24	3	US-09-117-860-26	Sequence 26, Appl
C 15	30.0	24	3	US-09-117-860-28	Sequence 28, Appl
C 16	30.0	24	3	US-09-117-860-30	Sequence 30, Appl
C 17	30.0	25	4	US-09-866-108A-4264	Sequence 4264, Ap
C 18	30.0	25	4	US-09-866-108A-4265	Sequence 4265, Ap
C 19	30.0	25	4	US-09-866-108A-5536	Sequence 5536, Ap
C 20	30.0	25	4	US-09-866-108A-5537	Sequence 5537, Ap
C 21	30.0	25	4	US-09-866-108A-5538	Sequence 5538, Ap
C 22	30.0	25	4	US-09-866-108A-5539	Sequence 5539, Ap
C 23	30.0	25	4	US-09-866-108A-5540	Sequence 5540, Ap
C 24	30.0	25	4	US-09-396-196G-43374	Sequence 43374, A
C 25	30.0	25	4	US-09-396-196G-43375	Sequence 43375, A
C 26	30.0	25	4	US-09-396-196G-82528	Sequence 82528, A
C 27	30.0	25	4	US-09-396-196G-99010	Sequence 99010, A
C 28	30.0	25	4	US-09-396-196G-99011	Sequence 99011, A
C 29	30.0	25	4	US-09-396-196G-104405	Sequence 104405, A
C 30	30.0	25	4	US-09-396-196G-109365	Sequence 109365, A
C 31	30.0	25	4	US-09-396-196G-127711	Sequence 127711, A
C 32	30.0	25	4	US-09-396-196G-127712	Sequence 127712, A
C 33	30.0	25	4	US-09-396-196G-127713	Sequence 127713, A
C 34	30.0	29	3	US-09-117-860-39	Sequence 39, Appl
C 35	30.0	30	2	US-08-961-749-3	Sequence 3, Appl
C 36	30.0	30	3	US-09-101-629A-22	Sequence 22, Appl
C 37	30.0	30	4	US-09-814-351-29	Sequence 29, Appl
C 38	30.0	33	3	US-08-554-385-27	Sequence 27, Appl
C 39	30.0	33	3	US-09-198-723A-73	Sequence 73, Appl
C 40	30.0	33	3	US-09-198-723A-74	Sequence 74, Appl
C 41	30.0	33	4	US-09-684-881-73	Sequence 73, Appl
C 42	30.0	33	4	US-09-684-881-74	Sequence 74, Appl
C 43	30.0	34	1	US-08-233-030-58	Sequence 58, Appl
C 44	30.0	34	2	US-08-577-492-15	Sequence 15, Appl
C 45	30.0	34	3	US-09-079-630-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-246-277A-3/c

; Sequence 3, Application US/09246277A

; Patent No. 6238866

; GENERAL INFORMATION:

; APPLICANT: Yeh, Homer R., Wick, Charles H.

; TITLE OF INVENTION: NOVEL DETECTOR FOR NUCLEIC ACID

; TYPING AND METHODS OF USING THE SAME

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Office of the Chief Counsel

; STREET: U.S. Army SBCCOM

; CITY: APG(EA)

; STATE: Maryland

; COUNTRY: United States of America

; ZIP: 21010-5423

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/246,277A

; FILING DATE: 08-Feb-1999

; CLASSIFICATION: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Biffoni, U. John

; REGISTRATION NUMBER: 39908

; REFERENCE/DOCKET NUMBER: 436-96

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (410) 436-1158

; TELEFAX: (410) 436-2534

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 32 base pairs

; TYPE: nucleic acid

```
/ STRANDEDNESS: not relevant
/ TOPOLOGY: not relevant
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Homo sapiens
/ POSITION IN GENOME:
/ CHROMOSOME/SEGMENT: 33.15 minisatellite
/ SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-246-277A-3
Alignment Scores:
Pred. No.: 367 Length: 32
Score: 33.00 Matches: 6
Percent Similarity: 66.67% Conservative: 0
Best Local Similarity: 66.67% Mismatches: 3
Query Match: 33.00% Indels: 0
DB: 3 Gaps: 0

US-09-632-036F-1 (1-19) x US-09-246-277A-3 (1-32)
Qy 9 LeuProAlaSerProGluThrHisLeu 17
Db 28 CTGCCACCTCTCCACCTGCCACCTC 2

RESULT 2
PCT-US93-07672-5
; Sequence 5, Application PC/TUS9307672
; GENERAL INFORMATION:
; APPLICANT: UNITED STATES OF AMERICA
; APPLICANT: as Represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: RECOMBINANT TOXIN WITH INCREASED
; TITLE OF INVENTION: HALF-LIFE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07672
; FILING DATE: 19930812
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/931,033
; FILING DATE: 14-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 15280-29
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
/ LOCATION: 1..36
/ OTHER INFORMATION: /function= "PCR primer"
/ OTHER INFORMATION: /product= "JGK-35"
/ OTHER INFORMATION: /note= "Used to amplify the CH3 domain."
PCT-US93-07672-5
Alignment Scores:
Pred. No.: 516 Length: 36
Score: 32.50 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 2
Query Match: 32.50% Indels: 3
DB: 5 Gaps: 1

US-09-632-036F-1 (1-19) x PCT-US93-07672-5 (1-36)
Qy 3 ThrAspMetLysLeuArgLeuProAlaSerProGluThrHis 16
Db 2 ACCTCTGTAAAGCTTCGC-----AGCCCCGAGAACCCAC 34

RESULT 3
US-09-663-834A-10/c
; Sequence 10, Application US/09663834A
; Patent No. 6613567
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF HER-2 EXPRESSION
; FILE REFERENCE: RTS-0033
; CURRENT APPLICATION NUMBER: US/09/663,834A
; CURRENT FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 48
; SEQ ID NO 10
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-663-834A-10
Alignment Scores:
Pred. No.: 276 Length: 18
Score: 32.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 32.00% Indels: 0
DB: 4 Gaps: 0

US-09-632-036F-1 (1-19) x US-09-663-834A-10 (1-18)
Qy 10 ProAlaSerProGluThr 15
Db 18 CCTGCCAGTCCCGAGACC 1

RESULT 4
US-09-396-196G-18778
; Sequence 18778, Application US/09396196G
; Patent No. 6821724
; GENERAL INFORMATION:
; APPLICANT: Michael Mittmann
; APPLICANT: David Mack
; APPLICANT: David Lockhart
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Methods of Genetic Analysis
; FILE REFERENCE: 3101.1
; CURRENT APPLICATION NUMBER: US/09/396,196G
; CURRENT FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: 60/100,678
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 127806
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18778
; LENGTH: 25
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2005, 10:12:52 ; Search time 268.171 Seconds
(without alignments)
419.415 Million cell updates/sec

Title: US-09-632-036F-1

Perfect score: 100

Sequence: 1 TGTDMKRLPASPETHLDM 19

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 3916100

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DSV=rlh
-Q=/cgn2_1/USPTO_spool_h/US09632036/runat_17062005_132235_8632/app_query.fasta_1.1791
-DB=N_Geneseq -QMT=Fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=40
-USER=US09632036 @CGN 1.1 1401 @runat_17062005_132235_8632 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37	37.0	35	12	ADL72974 Hepatitis
2	36	36.0	29	10	ACH01230 Stem loop
3	36	36.0	29	10	ACH01236 Stem loop
C 4	35	35.0	19	11	ADL79101 Human HER
5	35	35.0	19	11	ADL78852 Human HER

C	6	35	35.0	30	2	AAQ27073
C	7	35	35.0	32	12	ADN97675
C	8	35	35.0	37	9	ACC83681
C	9	34	34.0	33	10	AB225064
C	10	34	34.0	39	2	AAV48328
C	11	34	34.0	39	8	ACC42743
C	12	33	33.0	25	9	ACK30771
C	13	33	33.0	25	9	ACK11533
C	14	33	33.0	30	4	AAFP3070
C	15	33	33.0	32	5	AAQ04022
C	16	33	33.0	33	4	AAH75637
C	17	33	33.0	33	6	ABA01842
C	18	33	33.0	34	12	ACF36932
C	19	33	33.0	34	12	ACF36928
C	20	33	33.0	34	12	ACF36930
C	21	33	33.0	34	12	ACF36934
C	22	33	33.0	35	12	ADL72971
C	23	33	33.0	37	9	ACC83732
C	24	32.5	32.5	36	2	AAQ58499
C	25	32	32.0	18	6	AAQ38911
C	26	32	32.0	18	13	ADQ75952
C	27	32	32.0	18	13	ADQ75948
C	28	32	32.0	19	11	ADL78850
C	29	32	32.0	19	11	ADL79099
C	30	32	32.0	27	3	AAZ59942
C	31	32	32.0	30	4	AAH21866
C	32	32	32.0	32	2	AAT18055
C	33	32	32.0	33	2	AAQ52044
C	34	32	32.0	33	2	AAV72967
C	35	32	32.0	35	2	AAQ03931
C	36	32	32.0	35	4	AAI46388
C	37	32	32.0	37	2	AAQ91875
C	38	32	32.0	37	2	AAQ54779
C	39	32	32.0	37	3	AAZ92747
C	40	32	32.0	38	6	ABN84619
C	41	32	32.0	40	12	ADU53680
C	42	31	31.0	21	6	ABX09462
C	43	31	31.0	24	6	ABN84577
C	44	31	31.0	25	4	AAH39091
C	45	31	31.0	25	9	ACI23138

ALIGNMENTS

RESULT 1
ADL72974/C
ID ADL72974 standard; DNA; 35 BP.

XX

AC ADL72974;

XX

DT 17-JUN-2004 (first entry)

XX

DE Hepatitis C virus E1E2 polypeptide plasmid PCR primer #6.

XX

KW HCV; hepatitis; vaccine; hepatitis; infection; plasmid; ss; primer;

KM PCR.

XX

OS Hepatitis C virus.

OS Synthetic.

XX

PN WO2004024904-A2.

XX

PD 25-MAR-2004.

XX

PF 12-SEP-2003; 2003WO-IB003882.

XX

PR 13-SEP-2002; 2002EP-00292254.

PR 03-MAR-2003; 2003EP-00290505.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

PA Bartosch B, Cosset F;

XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 21, 2005, 10:28:03 ; Search time 2126.19 Seconds
(without alignments)
433.004 Million cell updates/sec

Title: US-09-632-036F-1
Perfect score: 100
Sequence: 1 TGTDMKRLPASPETHLDM 19

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 1692386

Minimum DB seq length: 0
Maximum DB seq length: 40

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPIO_spool_h/US09632036/runat_17062005_132237_8691/app_query.fasta_1.1791
-DB=GenEmbl -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=40
-USER=US09632036 @CGN 1 1 9703 @runat_17062005_132237_8691 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCKS=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
c 1	37	37.0	35	6	CQ785587 Sequence
c 2	37	37.0	35	6	CQ795329 Sequence
c 3	37	37.0	35	6	CQ867233 Sequence
c 4	35	35.0	30	9	X58166 Human T-cel

5	35	35.0	32	6	CQ801265 Sequence
c 6	33	33.0	30	6	AX107905 Sequence
c 7	33	33.0	30	6	AR154125 Sequence
8	33	33.0	35	6	CQ785584 Sequence
9	33	33.0	35	6	CQ795326 Sequence
10	33	33.0	35	6	CQ867230 Sequence
c 11	32	32.0	18	6	CQ841321 Sequence
c 12	32	32.0	18	6	CQ841325 Sequence
c 13	32	32.0	18	6	AR392095 Sequence
c 14	32	32.0	30	6	BD095106 N-Type ca
c 15	32	32.0	32	6	E10693 Primer. 9/1
16	32	32.0	33	6	I34919 Sequence 5
17	32	32.0	37	6	BD27235 Polyhydro
18	32	32.0	37	6	AR342651 Sequence
19	32	32.0	37	6	AR366501 Sequence
20	31	31.0	21	6	AX613664 Sequence
21	31	31.0	25	6	AX116764 Sequence
22	31	31.0	29	6	AX497614 Sequence
23	31	31.0	29	6	AX539926 Sequence
c 24	31	31.0	38	6	I58428 Sequence 12
25	31	31.0	38	6	AR200026 Sequence
26	31	31.0	38	6	AR227671 Sequence
27	31	31.0	39	10	MMIGH592 X77003 M.musculus
c 28	30	30.0	24	6	BD137624 Mutation
c 29	30	30.0	24	6	AR182613 Sequence
c 30	30	30.0	24	6	AR182615 Sequence
c 31	30	30.0	24	6	AR182617 Sequence
c 32	30	30.0	25	6	CQ619524 Sequence
c 33	30	30.0	25	6	CQ619525 Sequence
c 34	30	30.0	25	6	CQ620796 Sequence
c 35	30	30.0	25	6	CQ620797 Sequence
c 36	30	30.0	25	6	CQ620798 Sequence
c 37	30	30.0	25	6	CQ620799 Sequence
c 38	30	30.0	25	6	CQ620800 Sequence
c 39	30	30.0	25	6	AR460587 Sequence
c 40	30	30.0	25	6	AR460588 Sequence
c 41	30	30.0	25	6	AR461859 Sequence
c 42	30	30.0	25	6	AR461860 Sequence
c 43	30	30.0	25	6	AR461861 Sequence
c 44	30	30.0	25	6	AR461862 Sequence
c 45	30	30.0	25	6	AR461863 Sequence

ALIGNMENTS

RESULT 1	CQ785587/c	CQ785587	Sequence 7 from Patent EP1398371.	35 bp	DNA	linear	PAT 24-MAR-2004
LOCUS	CQ785587	CQ785587	CQ785587.1	GI:45720836			
DEFINITION	CQ785587	CQ785587	CQ785587.1	GI:45720836			
ACCESSION	CQ785587	CQ785587	CQ785587.1	GI:45720836			
VERSION	CQ785587	CQ785587	CQ785587.1	GI:45720836			
KEYWORDS	CQ785587	CQ785587	CQ785587.1	GI:45720836			
SOURCE	CQ785587	CQ785587	CQ785587.1	GI:45720836			
ORGANISM	CQ785587	CQ785587	CQ785587.1	GI:45720836			
REFERENCE	CQ785587	CQ785587	CQ785587.1	GI:45720836			
AUTHORS	CQ785587	CQ785587	CQ785587.1	GI:45720836			
TITLE	CQ785587	CQ785587	CQ785587.1	GI:45720836			
JOURNAL	CQ785587	CQ785587	CQ785587.1	GI:45720836			
FEATURES	CQ785587	CQ785587	CQ785587.1	GI:45720836			
source	CQ785587	CQ785587	CQ785587.1	GI:45720836			
ORIGIN	CQ785587	CQ785587	CQ785587.1	GI:45720836			
Alignment Scores:	CQ785587	CQ785587	CQ785587.1	GI:45720836			
Pred. No.:	CQ785587	CQ785587	CQ785587.1	GI:45720836			
Score:	CQ785587	CQ785587	CQ785587.1	GI:45720836			

Percent Similarity: 81.82% Conservative: 2
 Best Local Similarity: 63.64% Mismatches: 2
 Query Match: 37.00% Indels: 0
 DB: 6 Gaps: 0

US-09-632-036F-1 (1-19) x CQ785587 (1-35)

QY 7 LeuArgLeuProAlaSerProGluThrHisLeu 17
 ||| ::::::::::|
 DB 35 CTGACTGTGCCGCTTCAGCGAACCACCGTC 3

RESULT 2

CQ795329/c
 LOCUS CQ795329 35 bp DNA linear PAT 19-APR-2004

DEFINITION Sequence 7 from Patent WO2004024904.

ACCESSION CQ795329

VERSION CQ795329.1 GI:46407424

KEYWORDS synthetic construct

SOURCE other sequences; artificial sequences.

ORGANISM
 1
 REFERENCE
 AUTHORS Bartosch,B. and Cosset,F.L.
 TITLE Infectious hepatitisvirus pseudo-particles containing functional el,e2
 envelope proteins
 JOURNAL Patent: WO 2004024904-A 7 25-MAR-2004;
 INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE (INSERM)
 (FR)

FEATURES

source
 1..35
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="primer"

ORIGIN

Alignment Scores:
 Pred. No.: 1.18e+03 Length: 35
 Score: 37.00 Matches: 7
 Percent Similarity: 81.82% Conservative: 2
 Best Local Similarity: 63.64% Mismatches: 2
 Query Match: 37.00% Indels: 0
 DB: 6 Gaps: 0

US-09-632-036F-1 (1-19) x CQ795329 (1-35)

QY 7 LeuArgLeuProAlaSerProGluThrHisLeu 17
 ||| ::::::::::|
 DB 35 CTGACTGTGCCGCTTCAGCGAACCACCGTC 3

RESULT 3

CQ867233/c
 LOCUS CQ867233 35 bp DNA linear PAT 13-SEP-2004

DEFINITION Sequence 7 from Patent EP1454989.

ACCESSION CQ867233

VERSION CQ867233.1 GI:51997475

KEYWORDS synthetic construct

SOURCE other sequences; artificial sequences.

ORGANISM
 1
 REFERENCE
 AUTHORS Bartosch,B. and Cosset,F.L.
 TITLE Infectious HCV pseudo-particles containing native functional E1 and
 E2 envelope proteins
 JOURNAL Patent: EP 1454989-A 7 08-SEP-2004;
 INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE (INSERM)
 (FR)

FEATURES

source
 1..35
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="primer"

ORIGIN

Alignment Scores:
 Pred. No.: 1.18e+03 Length: 35
 Score: 37.00 Matches: 7
 Percent Similarity: 81.82% Conservative: 2
 Best Local Similarity: 63.64% Mismatches: 2
 Query Match: 37.00% Indels: 0
 DB: 6 Gaps: 0

US-09-632-036F-1 (1-19) x CQ867233 (1-35)

QY 7 LeuArgLeuProAlaSerProGluThrHisLeu 17
 ||| ::::::::::|
 DB 35 CTGACTGTGCCGCTTCAGCGAACCACCGTC 3

RESULT 4

HSTCELJ/c

LOCUS HSTCELJ 30 bp mRNA linear PRI 03-MAR-1999

DEFINITION Human T-cell receptor J-alpha region mRNA for J-alpha FR4.

ACCESSION X58166

VERSION X58166.1 GI:36782

KEYWORDS antigen receptor; J-alpha region; T-cell receptor; T-lymphocyte
 antigen receptor.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Pluschke,G., Ricken,G., Taube,H., Kroninger,S., Melchers,I.,
 Peter,H.H., Eichmann,K. and Krawinkel,U.
 TITLE Biased T cell receptor V alpha region repertoire in the synovial
 fluid of rheumatoid arthritis patients
 JOURNAL Eur. J. Immunol. 21 (11), 2749-2754 (1991)

MEDLINE 92037820

PUBMED 1657615

REFERENCE 2 (bases 1 to 30)

AUTHORS Krawinkel,U.

TITLE Direct Submission

JOURNAL Submitted (28-FEB-1991) U. Krawinkel, Klinische Forschergruppe f.
 Rheumatologie, Mooswaldallee 1-9, 7800 Freiburg, FRG
 See also X58157-X58168 for related sequences.

COMMENT

FEATURES

Location/Qualifiers

source

1..30

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/cell_line="synovial T-lymphocyte"

/tissue_type="lymphocyte"

gene

1..30

/gene="T-cell receptor J-alpha region"

misc_feature

1..30

/gene="T-cell receptor J-alpha region"

/product="V-alpha FR4"

/note="V region, FR4"

ORIGIN

Alignment Scores:
 Pred. No.: 2.17e+03 Length: 30
 Score: 35.00 Matches: 7
 Percent Similarity: 88.89% Conservative: 1
 Best Local Similarity: 77.78% Mismatches: 1
 Query Match: 35.00% Indels: 0
 DB: 9 Gaps: 0

US-09-632-036F-1 (1-19) x HSTCELJ (1-30)

QY 6 LysLeuArgLeuProAlaSerProGlu 14

||| ::::::::::|

DB 29 AAGGTGATTGTGCTTCTTCCCGAG 3

RESULT 5

CQ801265

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 3, 2005, 19:03:10 ; Search time 61.1644 Seconds
(without alignments)
120.143 Million cell updates/sec

Title: US-09-632-036F-1

Perfect score: 100

Sequence: 1 TCTDMKRLPASPETHLDM 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 888314

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A Geneseq 16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100	100.0	19	4	AAB68622 HER-2 B C
2	100	100.0	32	2	AAB59346 Aaw59346 Human Erb
3	100	100.0	32	2	AAY14575 Aay14575 Human erb
4	100	100.0	32	4	AAB62075 Aab62075 ErbB2 7C2
5	100	100.0	32	4	AAB48764 Aab48764 Human Erb
6	80	80.0	15	4	AAG89010 Aag89010 Her2/neu
7	80	80.0	15	4	AAG88392 Aag88392 HER2/NEU
8	79	79.0	15	3	AAY98884 Aay98884 HLA class
9	79	79.0	15	4	AAG89011 Aag89011 Her2/neu
10	79	79.0	15	4	AAG88332 Aag88332 HER2/NEU
11	78	78.0	25	4	AAG66872 Aag66872 Human Erb
12	73	73.0	15	3	AAY98840 Aay98840 HLA class
13	72	72.0	15	4	AAG66881 Aag66881 Human Erb
14	72	72.0	20	4	AAG66880 Aag66880 Human Erb
15	72	72.0	25	4	AAG66871 Aag66871 Human Erb
16	64	64.0	15	3	AAY98932 Aay98932 HLA class
17	64	64.0	15	4	AAG89066 Aag89066 HER2/neu
18	64	64.0	15	4	AAG88752 Aag88752 HER2/NEU
19	53	53.0	10	2	AAB61633 Aab61633 Peptide f
20	53	53.0	10	6	ABP74448 Abp74448 Human Her
21	53	53.0	10	7	ADC09307 Adc09307 Epitope w
22	53	53.0	10	8	ADE97864 Ade97864 Immunogen
23	51	51.0	10	6	ABP74442 Abp74442 Human Her
24	51	51.0	10	7	ADC09301 Adc09301 Epitope w
25	48	48.0	9	4	AAG88751 Aag88751 HER2/NEU

26	48	48.0	9	4	AAB68646 Aab68646 HER-2 CTL
27	48	48.0	9	6	ABP74447 Abp74447 Human Her
28	48	48.0	9	7	ADC09306 Adc09306 Epitope w
29	48	48.0	10	6	ABP74450 Abp74450 Human Her
30	48	48.0	10	7	ADC09309 Adc09309 Epitope w
31	47	47.0	10	4	AAG88919 Aag88919 HER2/neu
32	45	45.0	9	4	AAG88391 Aag88391 HER2/NEU
33	45	45.0	9	4	AAG88331 Aag88331 HER2/NEU
34	45	45.0	15	4	AAG66884 Aag66884 Human Erb
35	45	45.0	20	4	AAG66882 Aag66882 Human Erb
36	45	45.0	25	4	AAG66885 Aag66885 Human Erb
37	44	44.0	8	6	ABP74446 Abp74446 Human Her
38	44	44.0	8	7	ADC09305 Adc09305 Epitope w
39	44	44.0	9	6	ABP74449 Abp74449 Human Her
40	44	44.0	9	7	ADC09308 Adc09308 Epitope w
41	40	40.0	7	8	ADP80317 Adp80317 Human HLA
42	40	40.0	8	6	ABP74441 Abp74441 Human Her
43	40	40.0	8	7	ADC09300 Adc09300 Epitope w
44	40	40.0	8	8	ADN64551 Adn64551 HLA bindi
45	40	40.0	8	8	ADP80286 Adp80286 Human HLA
46	40	40.0	11	8	ADN64641 Adn64641 HLA bindi
47	40	40.0	11	8	ADP80287 Adp80287 Human HLA
48	40	40.0	18	4	ABB41449 Abb41449 Peptide #
49	40	40.0	18	4	AAM62323 Aam62323 Human bra
50	38	38.0	33	8	ADS84891 Ads84891 Beta-eite
51	37	37.0	14	4	AAM53151 Aam53151 Human con
52	36	36.0	31	4	AUU31251 Auu31251 Novel hum
53	35	35.0	36	8	ADJ25789 Adj25789 MDM2 bind
54	34	34.0	21	4	AAM20779 Aam20779 Peptide #
55	34	34.0	21	4	ABB42333 Abb42333 Peptide #
56	34	34.0	21	4	AAM36139 Aam36139 Peptide #
57	34	34.0	21	4	ABM25824 Abm25824 Protein #
58	34	34.0	21	4	AAM76032 Aam76032 Human bon
59	34	34.0	21	4	AAM63219 Aam63219 Human bra
60	34	34.0	21	4	ABG57758 Abg57758 Human liv
61	34	34.0	21	5	ABG45430 Abg45430 Human pep
62	34	34.0	27	5	AUU89413 Auu89413 Insulin/i
63	34	34.0	33	6	ABU08630 Abu08630 LSI47-enc
64	34	34.0	37	8	ABO53961 Abo53961 Human gen
65	33	33.0	10	6	ABF74445 Abf74445 Human Her
66	33	33.0	10	7	ADC09304 Adc09304 Epitope w
67	33	33.0	11	4	AAG88920 Aag88920 HER2/neu
68	33	33.0	15	6	ADA00906 Ada00906 Mouse GL
69	33	33.0	24	7	ADC27958 Adc27958 Synthetic
70	33	33.0	29	8	ABO57829 Abo57829 Human gen
71	33	33.0	30	5	ABG62068 Abg62068 Human pro
72	33	33.0	37	8	AAO09019 Aao09019 Human pol
73	33	33.0	37	8	ADJ45726 Adj45726 Signal pe
74	33	33.0	38	2	AAB58380 Aab58380 TSAR bind
75	32	32.0	9	2	AAY37971 Aay37971 Human CER
76	32	32.0	9	2	AAY45541 Aay45541 Immunogen
77	32	32.0	14	4	AAM53150 Aam53150 Human con
78	32	32.0	15	2	AAR97899 Aar97899 Japan ced
79	32	32.0	16	5	ABB06398 Abb06398 Intracell
80	32	32.0	24	7	AAY58888 Aay58888 Polyketid
81	32	32.0	24	7	ADC27959 Adc27959 Synthetic
82	32	32.0	26	2	AAY12216 Aay12216 Human 5'
83	32	32.0	27	8	ADP81213 Adp81213 Protein o
84	32	32.0	32	3	ADQ95637 Adq95637 Aggrecon
85	32	32.0	34	3	AAB30463 Aab30463 Amino aci
86	32	32.0	35	4	AAM15419 Aam15419 Peptide #
87	32	32.0	35	4	ABB34431 Abb34431 Peptide #
88	32	32.0	35	4	AAM27909 Aam27909 Peptide #
89	32	32.0	35	4	ABB19841 Abb19841 Protein #
90	32	32.0	35	4	AAM67613 Aam67613 Human bon
91	32	32.0	35	4	AAM55217 Aam55217 Human bra
92	32	32.0	35	4	AAM03181 Aam03181 Peptide #
93	32	32.0	39	5	ABB06391 Abb06391 Human int
94	32	32.0	39	5	ABB06392 Abb06392 Carp intr
95	32	32.0	39	5	ABB06406 Abb06406 Intracell
96	32	32.0	40	3	AAG33986 Aag33986 Arabidops
97	32	32.0	40	5	ABB06397 Abb06397 Intracell
98	31	31.0	12	2	AAR23579 Aar23579 HTLV-II d

99 31 31.0 13 8 ADM12693 Adm12693 Ii key/WH
100 31 31.0 13 8 ADO38929 ADO38929 Anthrax t

ALIGNMENTS

```

RESULT 1
AAB68622
ID AAB68622 standard; peptide; 19 AA.
XX
AC AAB68622;
XX
DT 30-APR-2001 (first entry)
XX
DE HER-2 B cell epitope #1.
XX
KW Cytostatic; immune response; HER-2; human; epitope; cancer; breast;
KW ovarian; lung; prostate; colon.
XX
OS Homo sapiens.
XX
PN WO200108636-A2.
XX
PD 08-FEB-2001.
XX
PF 03-AUG-2000; 2000WO-US021222.
XX
PR 03-AUG-1999; 99US-0146869P.
XX
PA (OHIS ) UNIV OHIO STATE.
XX
PI Kaumaya PT, Stevens VC, Triozzi PL;
XX
DR WPI; 2001-182849/18.
XX
PT Compositions comprising polypeptides and polynucleotides for stimulating
PT the immune system and for treating malignancies associated with
PT overexpression of the HER-2 protein.
XX
PS Claim 1; Page 37; 51pp; English.
XX
CC The present invention relates to compositions for stimulating the immune
CC system and for treating malignancies associated with overexpression of
CC the HER-2 protein. The compositions comprise immunogenic groups of the
CC HER-2 proteins. The present sequence is one such peptide used in the
CC compositions of the present invention. The compositions can be used for
CC treating cancer, e.g. breast, ovarian, lung, prostate and colon cancers
XX
SQ Sequence 19 AA;

Query Match 100.0%; Score 100; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTDMKRLRPASPETHLDM 19
DB 1 TGTDMKRLRPASPETHLDM 19

RESULT 2
AAW59346
ID AAW59346 standard; protein; 32 AA.
XX
AC AAW59346;
XX
DT 08-SEP-1998 (first entry)
XX
DE Human ErbB2 extracellular domain protein fragment (aa 22-53).
XX
KW ErbB2; epitope; monoclonal antibody; identification; apoptosis;
KW detection; treatment; disorder; tumour; benign; malignant; leukaemia;
KW lymphoid malignancy; inflammation; angiogenic; immunological.
XX

```

```

XX Homo sapiens.
XX WO9817797-A1.
XX 30-APR-1998.
XX 09-OCT-1997; 97WO-US018385.
XX 18-OCT-1996; 96US-00731794.
XX (GETH ) GENENTECH INC.
XX (TEXA ) UNIV TEXAS SYSTEM.
XX Fendly BM, Phillips GD, Scheuermann RH, Uhr JW;
XX WPI; 1998-261493/23.
XX New anti-ErbB2 antibodies - which induce apoptosis in cells which
XX overexpress ErbB2, used to treat e.g. tumours or inflammatory, angiogenic
XX or immunologic disorders.
XX Disclosure; Page 45; 72pp; English.
XX This sequence represents a fragment of the human ErbB2 extracellular
XX domain corresponding to amino acids 22-53 which is capable of binding the
XX monoclonal antibodies 7C3 and 7F3. This region is used in a method to
XX identify antibodies which bind to ErbB2 and induce apoptosis of a cell
XX which overexpresses ErbB2. The antibodies can also be used to detect
XX ErbB2 on a cell. The antibodies can be used for treating disorders such
XX as benign or malignant tumours, (e.g. renal, liver, kidney, bladder,
XX breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
XX thyroid, hepatic carcinomas, sarcomas, glioblastomas, and various head
XX and neck tumours), leukemia's and lymphoid malignancies, other disorders
XX such as neuronal, glial, astrocytal, hypochalamic and other glandular,
XX macrophagal, epithelial, stromal and blastocoelec disorders, and
XX inflammatory, angiogenic and immunologic disorders
XX Sequence 32 AA;

Query Match 100.0%; Score 100; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 3.3e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGTDMKRLRPASPETHLDM 19
DB 6 TGTDMKRLRPASPETHLDM 24

RESULT 3
AAV14575
ID AAV14575 standard; peptide; 32 AA.
XX
AC AAV14575;
XX
DT 14-SEP-1999 (first entry)
XX
DE Human erbB2 receptor-derived epitope 7C2/7F3.
XX
KW Domain; human; erbB2; receptor; antibody; overexpression; anthracycline;
KW chemotherapeutic agent; tumour; cancer; breast; squamous cell; cervical;
KW non-small-cell lung; gastrointestinal; pancreatic; glioblastoma; ovarian;
KW liver; bladder; hepatoma; colon; colorectal; endometrial carcinoma;
KW salivary gland; kidney; prostate; vulva; thyroid; hepatic carcinoma;
KW head-and-neck; epitope.
XX
OS Homo sapiens.
XX
PN WO9311140-A1.
XX
PD 24-JUN-1999.
XX
PF 10-DEC-1998; 98WO-US026266.
XX

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 18, 2005, 00:00:13 ; Search time 1876.93 Seconds
(without alignments)
385.322 Million cell updates/sec

Title: US-09-632-036F-1

Perfect score: 100

Sequence: 1 TGTDMKRLPASPETHLDM 19

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 94960

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool_P/US09632036/runat.16062005.151337.11030/app_query.fasta.1.1791
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=40
-USPR=US09632036 @CGN 1.1 16765 @runat.16062005.151337.11030 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gest1:*
9: gb_gest2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	36.0	40	2 BE733350	601570571
2	35	35.0	39	9 AL943196	Arabidops
3	30	30.0	28	1 AI769631	AI769631 wj25a04.x
4	30	30.0	31	1 AI153615	AI153615 vz89f11.x
5	30	30.0	34	1 AA912855	AA912855 ol32d01.s
6	30	30.0	39	8 BZ664670	BZ664670 SALK_0997
7	30	30.0	40	1 AI155888	AI155888 ue07c10.x
8	29	29.0	28	1 AI441029	AI441029 sa58e02.y
9	29	29.0	31	8 AZ481183	AZ481183 IM0303G17

c	10	29	29.0	31	8	AZ648445	IM0517A02
c	11	29	29.0	34	8	BH907784	SALK 0440
c	12	29	29.0	37	7	C0793217	NT017A.H0
c	13	29	29.0	37	8	AZ662400	IM0541A02
c	14	29	29.0	40	8	BZ355646	SALK 1271
c	15	28	28.0	30	9	AL769095	Arabidops
c	16	28	28.0	33	9	TA318G05P	AL492869 T. brucei
c	17	28	28.0	34	1	AA212311	mu43ell1.r
c	18	28	28.0	34	1	AA259780	va87b10.r
c	19	28	28.0	37	8	BH847217	SALK 0446
c	20	28	28.0	37	9	CG730937	1115130H0
c	21	28	28.0	38	4	BI547045	603190269
c	22	28	28.0	40	8	AZ622990	IM0460M12
c	23	27	27.0	27	8	AZ659769	IM0537119
c	24	27	27.0	30	8	BH740820	KG04070-5
c	25	27	27.0	31	9	AL752303	Arabidops
c	26	27	27.0	32	8	AZ351426	IM0089805
c	27	27	27.0	33	8	AZ445447	IM0241G16
c	28	27	27.0	34	1	AA282064	AZ282064 zt02e01.s
c	29	27	27.0	34	8	AZ616433	IM0446E15
c	30	27	27.0	34	9	TA127C11P	AL465493 T. brucei
c	31	27	27.0	36	2	BE385175	601274641
c	32	27	27.0	37	1	AA897070	al06h07.s
c	33	27	27.0	37	6	CA587203	LBEO9p63
c	34	27	27.0	39	2	BE739441	601556230
c	35	27	27.0	40	1	AA069196	zm11b11.r
c	36	27	27.0	40	8	BH813908	SALK 0654
c	37	27	27.0	40	9	CG784612	RRR688 Ba
c	38	26.5	26.5	40	1	AI421705	tf54e02.x
c	39	26	26.0	20	4	BM393606	50072-2-1
c	40	26	26.0	23	8	AZ840388	2M0136J24
c	41	26	26.0	25	8	AZ506197	IM0347C11
c	42	26	26.0	25	8	AZ646723	AZ646723 IM0512F19
c	43	26	26.0	29	8	AZ596214	IM0409A21
c	44	26	26.0	29	9	TA246H09Q	AL488820 T. brucei
c	45	26	26.0	31	1	AL044950	DKF2p434P

ALIGNMENTS

RESULT 1
BE733350
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

601570571F1 NIH_MGC_21 Homo sapiens linear EST 15-SEP-2000
40 bp mRNA
mRNA sequence.
BE733350
BE733350.1 GI:10147342
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 40)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LUCM342 row: n column: 17.

FEATURES
source

Location/Qualifiers
1..40
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3845200"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"

/clone lib="NIH MGC 21"
 /note="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
 Pred. No.: 8.49e+03 Length: 40
 Score: 36.00 Matches: 8
 Percent Similarity: 72.73% Conservative: 0
 Best Local Similarity: 72.73% Mismatches: 3
 Query Match: 36.00% Indels: 0
 DB: 2 Gaps: 0

US-09-632-036F-1 (1-19) x BE733350 (1-40)

QY 3 ThrAspMetLysLeuArgLeuProLaserPro 13
 DB 6 ACAGACAGGAACACTACGGCTAGTGCAGGCCG 38

RESULT 2

AL943196/c
 LOCUS AL943196 39 bp DNA linear GSS 01-APR-2004
 DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-272D12-015091,
 genomic survey sequence.

ACCESSION AL943196
 VERSION AL943196.1 GI:24399794
 KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 Li,Y., Rosso,M.G., Strizhov,N., Viehoveer,P. and Weishaar,B.
 GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for
 the identification of T-DNA insertion mutants in Arabidopsis
 thaliana

Bioinformatics 19 (11), 1441-1442 (2003)

JOURNAL MEDLINE
 PUBMED 22755829
 12874060

REFERENCE 2 Rosso,M.G., Li,Y., Strizhov,N., Reiss,B., Dekker,K. and
 Weishaar,B.

TITLE An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for
 flanking sequence tag-based reverse genetics

JOURNAL Plant Mol. Biol. 53 (1-2), 247-259 (2003)
 MEDLINE 23117147
 PUBMED 14756321

REFERENCE 3 Strizhov,N., Li,Y., Rosso,M.G., Viehoveer,P., Dekker,K.A. and
 Weishaar,B.

TITLE High-throughput generation of sequence indexes from T-DNA
 mutagenized Arabidopsis thaliana lines

Biotechniques 35 (6), 1164-1168 (2003)

JOURNAL PUBMED 14682050

REFERENCE 4 (bases 1 to 39)
 Rosso,M.G., Li,Y., Strizhov,N. and Weishaar,B.

TITLE Direct Submission
 Submitted (31-MAR-2004) Weishaar B., Max-Planck-Institut fuer
 Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

JOURNAL This sequence has been recovered from the left border of the T-DNA.
 It indicates an insertion within the locus defined by BAC clone
 MDN11. Details on the protocols used for generation of the sequence
 are described in References 1-3. The sequences are generated at the
 MPI for Plant Breeding Research in the context of the GABI-Kat
 project. GABI-Kat is part of the German Plant Genomics program
 designated 'GABI'. Information on line availability can be found

at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

Location/Qualifiers
 1..39
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clones="GK-272D12-015091"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 /ecotype="Col-0"
 /note="PCR was performed on DNA from Arabidopsis thaliana
 plants (Ti) which were transformed with the T-DNA from
 vector pAC161 (GenBank accession number: AJ537514). The
 lines contain one or more T-DNA insertions. The DNA
 fragment(s) resulting from the PCR were directly sequenced
 to determine the genomic sequence flanking the insertion.
 T-DNA derived sequences were removed."

ORIGIN

Alignment Scores:
 Pred. No.: 1.2e+04 Length: 39
 Score: 35.00 Matches: 6
 Percent Similarity: 77.78% Conservative: 1
 Best Local Similarity: 66.67% Mismatches: 2
 Query Match: 35.00% Indels: 0
 DB: 9 Gaps: 0

US-09-632-036F-1 (1-19) x AL943196 (1-39)

QY 8 ArgLeuProAlaSerProGluThrHis 16

DB 35 AGACTTCCTCCGCTAACCCAGCGAGCAT 9

RESULT 3

AI769631/c

LOCUS AI769631

DEFINITION wJ25a04.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE:2403822 3',
 similar to TR:Q92465 Q92466 DDBB P48. ;, mRNA sequence.

ACCESSION AI769631
 VERSION AI769631.1 GI:5236140
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 28)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Insert Length: 1994 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers
 1..28

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2403822"

/tissue_type="2 pooled tumors (clear cell type)"

/lab_host="DH10B"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 18, 2005, 04:41:09 ; Search time 282.195 Seconds
(without alignments)
417.950 Million cell updates/sec

Title: US-09-632-036F-1

Perfect score: 100

Sequence: 1 TGTDMKLRLPASPETHLDM 19

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6054689 seqs, 3103772919 residues

Total number of hits satisfying chosen parameters: 5878042

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USFPO_spool_p/US09632036/runat_16062005_151339_11129/app_query.fasta_1.1791
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=1 -MATRIX=bl6sum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=40 -USER=US09632036 @CNC 1.1 1967 @runat_16062005_151339_11129 -NCFU=6
-ICFU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/1/pubpna/US10J_PUBCOMB.seq:
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:
26: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
27: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	35.0	19	16	US-10-251-117-17
c 2	35	35.0	19	16	US-10-251-117-266
c 3	35	35.0	25	21	US-10-719-900-80309
c 4	35	35.0	25	21	US-10-719-900-212555
c 5	35	35.0	25	21	US-10-719-900-249478
c 6	35	35.0	25	21	US-10-719-900-346374
c 7	34	34.0	25	21	US-10-719-900-305176
c 8	34	34.0	39	14	US-10-057-136-15
c 9	34	34.0	39	16	US-10-265-811-9
c 10	33	33.0	25	15	US-10-098-263B-111514
c 11	33	33.0	25	15	US-10-098-263B-130752
c 12	33	33.0	25	21	US-10-719-900-196604
c 13	33	33.0	25	21	US-10-719-900-244731
c 14	33	33.0	25	21	US-10-719-900-395120
c 15	33	33.0	25	21	US-10-719-900-812465
c 16	33	33.0	25	21	US-10-719-900-943313
c 17	33	33.0	25	21	US-10-719-900-943314
c 18	33	33.0	25	21	US-10-719-900-966842
c 19	33	33.0	25	21	US-10-956-157-178532
c 20	33	33.0	30	15	US-10-134-643-21
c 21	32	32.0	19	16	US-10-251-117-15
c 22	32	32.0	19	16	US-10-251-117-264
c 23	32	32.0	25	21	US-10-719-900-44912
c 24	32	32.0	25	21	US-10-719-900-218229
c 25	32	32.0	25	21	US-10-719-900-218230
c 26	32	32.0	25	21	US-10-719-900-370643
c 27	32	32.0	25	21	US-10-719-900-542722
c 28	32	32.0	25	21	US-10-719-900-615040
c 29	32	32.0	25	21	US-10-809-189-18778
c 30	32	32.0	27	14	US-10-183-728-4
c 31	32	32.0	35	15	US-10-203-821-10
c 32	32	32.0	37	9	US-09-944-243-4
c 33	32	32.0	37	17	US-10-444-031-4
c 34	32	32.0	38	9	US-09-881-823-24
c 35	32	32.0	40	17	US-10-461-790-74
c 36	31	31.0	24	14	US-10-024-648-2
c 37	31	31.0	25	15	US-10-098-263B-23129
c 38	31	31.0	25	21	US-10-719-900-67215
c 39	31	31.0	25	21	US-10-719-900-99561
c 40	31	31.0	25	21	US-10-719-900-196313
c 41	31	31.0	25	21	US-10-719-900-196314
c 42	31	31.0	25	21	US-10-719-900-212556
c 43	31	31.0	25	21	US-10-719-900-279126
c 44	31	31.0	25	21	US-10-719-900-299529
c 45	31	31.0	25	21	US-10-719-900-343564

ALIGNMENTS

RESULT 1
US-10-251-117-17
; Sequence 17, Application US/102511117
; Publication No. US20030170891A1
; GENERAL INFORMATION:
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: RNA Interference Mediated Inhibition of Epidermal Growth Factor R
; FILE REFERENCE: 900/042 (MBH02-468-A)
; CURRENT APPLICATION NUMBER: US/10/251,117
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US 60/393,924
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 10/163,552
; PRIOR FILING DATE: 2002-06-06

```
; PRIOR APPLICATION NUMBER: US 60/358,580
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 09/916,466
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/296,249
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 1213
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Target sequence/siNA sense
US-10-251-117-17
Alignment Scores:
Pred. No.: 394 Length: 19
Score: 35.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.00% Indels: 0
DB: 16 Gaps: 0
US-09-632-036F-1 (1-19) x US-10-251-117-17 (1-19)
QY 13 ProGluThrHisLeuAsp 18
DB 1 CCCGAGACCCACCCUGGAC 18
RESULT 2
US-10-251-117-266/c
; Sequence 266, Application US/10251117
; Publication No. US20030170891A1
; GENERAL INFORMATION:
; APPLICANT: McSwiggen Pharmaceuticals, Inc.
; TITLE OF INVENTION: RNA Interference Mediated Inhibition of Epidermal Growth Factor R
; FILE REFERENCE: 900/042 (MHB02-468-A)
; CURRENT APPLICATION NUMBER: US/10/251,117
; CURRENT FILING DATE: 2003-02-24
; PRIOR APPLICATION NUMBER: US 60/393,924
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 10/163,552
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/358,580
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 09/916,466
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/296,249
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 1213
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 266
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: siNA antisense region
US-10-251-117-266
Alignment Scores:
Pred. No.: 394 Length: 19
Score: 35.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.00% Indels: 0
DB: 16 Gaps: 0
US-09-632-036F-1 (1-19) x US-10-251-117-266 (1-19)
QY 13 ProGluThrHisLeuAsp 18
; PRIOR APPLICATION NUMBER: US 60/358,580
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: US 09/916,466
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/296,249
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 1213
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Target sequence/siNA sense
US-10-251-117-17
Alignment Scores:
Pred. No.: 394 Length: 19
Score: 35.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 35.00% Indels: 0
DB: 16 Gaps: 0
US-09-632-036F-1 (1-19) x US-10-251-117-17 (1-19)
QY 13 ProGluThrHisLeuAsp 18
DB 1 CCCGAGACCCACCCUGGAC 18
RESULT 3
US-10-719-900-80309/c
; Sequence 80309, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 80309
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; OTHER INFORMATION: Description of Artificial Sequence: Target sequence/siNA sense
US-10-719-900-80309
Alignment Scores:
Pred. No.: 525 Length: 25
Score: 35.00 Matches: 5
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 62.50% Mismatches: 0
Query Match: 35.00% Indels: 0
DB: 21 Gaps: 0
US-09-632-036F-1 (1-19) x US-10-719-900-80309 (1-25)
QY 9 LeuProAlaSerProGluThrHis 16
DB 25 TTGCCGAGCAATCCACAACTCAT 2
RESULT 4
US-10-719-900-212555
; Sequence 212555, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002-11-20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 212555
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
; OTHER INFORMATION: Description of Artificial Sequence: Target sequence/siNA sense
US-10-719-900-212555
Alignment Scores:
Pred. No.: 525 Length: 25
Score: 35.00 Matches: 6
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 85.71% Mismatches: 0
Query Match: 35.00% Indels: 0
DB: 21 Gaps: 0
US-09-632-036F-1 (1-19) x US-10-719-900-212555 (1-25)
QY 12 SerProGluThrHisLeuAsp 18
DB 3 TCTCCGAGACCCATTGGAC 23
RESULT 5
US-10-719-900-249478
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 18, 2005, 00:16:03 ; Search time 58.719 Seconds
(without alignments)
529.458 Million cell updates/sec

Title: US-09-632-036F-1

Perfect score: 100

Sequence: 1 TGTDMKRLRASPETHLDM 19

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 1132692

Minimum DB seq length: 0

Maximum DB seq length: 40

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool_p/US09632036/runat_16062005_151338_11045/app_query.fasta_1.1791
-DB=Issued_Patents_NA -QWTF=fastcap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=40
-USER=US09632036 -CGEN 1 1 392 @runat_16062005_151338_11045 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	33	33.0	32	3	US-09-246-277A-3
C 2	32.5	32.5	36	5	PCT-US93-07672-5
C 3	32	32.0	18	4	US-09-663-834A-10
4	32	32.0	25	4	US-09-396-196G-18778
5	32	32.0	33	1	US-08-435-350-5
6	32	32.0	37	3	US-09-366-920-4
7	32	32.0	37	4	US-09-944-243-4
8	31	31.0	25	4	US-09-396-196G-16697
9	31	31.0	25	4	US-09-396-196G-61016
C 10	31	31.0	25	4	US-09-396-196G-109354
C 11	31	31.0	38	1	US-08-360-841-12
C 12	31	31.0	38	3	US-09-308-160B-8
					Sequence 3, Appli
					Sequence 5, Appli
					Sequence 10, Appl
					Sequence 18778, A
					Sequence 5, Appli
					Sequence 4, Appli
					Sequence 16697, A
					Sequence 61016, A
					Sequence 109354,
					Sequence 12, Appl
					Sequence 8, Appli

13	31	31.0	38	3	US-09-456-399-8	Sequence 8, Appli
C 14	30	30.0	24	3	US-09-117-860-26	Sequence 28, Appl
C 15	30	30.0	24	3	US-09-117-860-28	Sequence 26, Appl
C 16	30	30.0	24	3	US-09-117-860-30	Sequence 30, Appl
C 17	30	30.0	25	4	US-09-866-108A-4264	Sequence 4264, Ap
C 18	30	30.0	25	4	US-09-866-108A-4265	Sequence 4265, Ap
C 19	30	30.0	25	4	US-09-866-108A-5536	Sequence 5536, Ap
C 20	30	30.0	25	4	US-09-866-108A-5537	Sequence 5537, Ap
C 21	30	30.0	25	4	US-09-866-108A-5538	Sequence 5538, Ap
C 22	30	30.0	25	4	US-09-866-108A-5539	Sequence 5539, Ap
C 23	30	30.0	25	4	US-09-866-108A-5540	Sequence 5540, Ap
C 24	30	30.0	25	4	US-09-396-196G-43374	Sequence 43374, A
25	30	30.0	25	4	US-09-396-196G-43375	Sequence 43375, A
26	30	30.0	25	4	US-09-396-196G-82528	Sequence 82528, A
27	30	30.0	25	4	US-09-396-196G-99010	Sequence 99010, A
28	30	30.0	25	4	US-09-396-196G-99011	Sequence 99011, A
29	30	30.0	25	4	US-09-396-196G-104405	Sequence 104405,
C 30	30	30.0	25	4	US-09-396-196G-109365	Sequence 109365,
C 31	30	30.0	25	4	US-09-396-196G-127711	Sequence 127711,
C 32	30	30.0	25	4	US-09-396-196G-127712	Sequence 127712,
C 33	30	30.0	25	4	US-09-396-196G-127713	Sequence 127713,
C 34	30	30.0	29	3	US-09-117-860-39	Sequence 39, Appl
C 35	30	30.0	30	2	US-08-961-749-3	Sequence 3, Appli
C 36	30	30.0	30	3	US-09-101-629A-22	Sequence 22, Appl
37	30	30.0	30	4	US-09-814-351-29	Sequence 29, Appl
C 38	30	30.0	33	3	US-08-554-385-27	Sequence 27, Appl
C 39	30	30.0	33	3	US-09-198-723A-73	Sequence 73, Appl
C 40	30	30.0	33	3	US-09-198-723A-74	Sequence 74, Appl
C 41	30	30.0	33	4	US-09-684-881-73	Sequence 73, Appl
C 42	30	30.0	33	4	US-09-684-881-74	Sequence 74, Appl
C 43	30	30.0	34	1	US-08-233-030-58	Sequence 58, Appl
C 44	30	30.0	34	2	US-08-577-492-15	Sequence 15, Appl
C 45	30	30.0	34	3	US-09-079-630-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-246-277A-3/c
; Sequence 3, Application US/09246277A
; Patent No. 6238866
; GENERAL INFORMATION:
; APPLICANT: Yeh, Homer R., Wick, Charles H.
; TITLE OF INVENTION: NOVEL DETECTOR FOR NUCLEIC ACID
; TYPING AND METHODS OF USING THE SAME
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Office of the Chief Counsel
; STREET: U.S. Army SBCCOM
; CITY: APG(IEA)
; STATE: Maryland
; COUNTRY: United States of America
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/246.277A
; FILING DATE: 08-Feb-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Biffoni, U. John
; REGISTRATION NUMBER: 39908
; REFERENCE/DOCKET NUMBER: 436-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (410) 436-1158
; TELEFAX: (410) 436-2534
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 base pairs
; TYPE: nucleic acid

```
/ ; STRANDEDNESS: not relevant
/ ; TOPOLOGY: not relevant
/ ; MOLECULE TYPE: DNA (genomic)
/ ; HYPOTHETICAL: NO
/ ; ANTI-SENSE: NO
/ ; ORIGINAL SOURCE:
/ ; ORGANISM: Homo sapiens
/ ; POSITION IN GENOME:
/ ; CHROMOSOME/SEGMENT: 33.15 minisatellite
/ ; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-246-277A-3
Alignment Scores:
Pred. No.: 367 Length: 32
Score: 33.00 Matches: 6
Percent Similarity: 66.67% Conservative: 0
Best Local Similarity: 66.67% Mismatches: 3
Query Match: 33.00% Indels: 0
DB: 3 Gaps: 0
US-09-632-036F-1 (1-19) x US-09-246-277A-3 (1-32)
QY 9 LeuProAlaSerProGluThrHisLeu 17
Db 28 CTGCCACCTCTCCACCTGCCACCTC 2
RESULT 2
PCT-US93-07672-5
/ ; SEQUENCE 5, Application PC/TUS9307672
/ ; GENERAL INFORMATION:
/ ; APPLICANT: UNITED STATES OF AMERICA
/ ; APPLICANT: as Represented by the Secretary of the
/ ; APPLICANT: Department of Health and Human Services
/ ; TITLE OF INVENTION: RECOMBINANT TOXIN WITH INCREASED
/ ; TITLE OF INVENTION: HALF-LIFE
/ ; NUMBER OF SEQUENCES: 8
/ ; CORRESPONDENCE ADDRESS:
/ ; ADDRESSER: Townsend and Townsend Kourie and Crew
/ ; STREET: Steuart Street Tower, One Market Plaza
/ ; CITY: San Francisco
/ ; STATE: California
/ ; COUNTRY: US
/ ; ZIP: 94105-1492
/ ; COMPUTER READABLE FORM:
/ ; MEDIUM TYPE: Floppy disk
/ ; COMPUTER: IBM PC compatible
/ ; OPERATING SYSTEM: PC-DOS/MS-DOS
/ ; SOFTWARE: PatentIn Release #1.0, Version #1.25
/ ; CURRENT APPLICATION DATA:
/ ; APPLICATION NUMBER: PCT/US93/07672
/ ; FILING DATE: 19930812
/ ; CLASSIFICATION:
/ ; PRIOR APPLICATION DATA:
/ ; APPLICATION NUMBER: US 07/931,033
/ ; FILING DATE: 14-AUG-1992
/ ; ATTORNEY/AGENT INFORMATION:
/ ; NAME: Weber, Ellen L.
/ ; REGISTRATION NUMBER: 32,762
/ ; REFERENCE/DOCKET NUMBER: 15280-29
/ ; TELECOMMUNICATION INFORMATION:
/ ; TELEPHONE: (415) 543-9600
/ ; TELEFAX: (415) 543-5043
/ ; INFORMATION FOR SEQ ID NO: 5:
/ ; SEQUENCE CHARACTERISTICS:
/ ; LENGTH: 36 base pairs
/ ; TYPE: nucleic acid
/ ; STRANDEDNESS: single
/ ; TOPOLOGY: linear
/ ; MOLECULE TYPE: DNA (genomic)
/ ; ORIGINAL SOURCE:
/ ; ORGANISM: Homo sapiens
/ ; FEATURE:
/ ; NAME/KEY: misc_feature
```

```
/ ; LOCATION: 1..36 /function= "PCR primer"
/ ; OTHER INFORMATION: /product= "JGK-35"
/ ; OTHER INFORMATION: /note= "Used to amplify the CH3 domain."
PCT-US93-07672-5
Alignment Scores:
Pred. No.: 516 Length: 36
Score: 32.50 Matches: 8
Percent Similarity: 64.29% Conservative: 1
Best Local Similarity: 57.14% Mismatches: 2
Query Match: 32.50% Indels: 3
DB: 5 Gaps: 1
US-09-632-036F-1 (1-19) x PCT-US93-07672-5 (1-36)
QY 3 ThrAspMetLysLeuArgLeuProAlaSerProGluThrHis 16
Db 2 ACCTCTGTAAAGCTTCGC-----AGCCCGAGAGACCAC 34
RESULT 3
US-09-663-834A-10/C
/ ; Sequence 10, Application US/09663834A
/ ; Patent No. 6613567
/ ; GENERAL INFORMATION:
/ ; APPLICANT: C. Frank Bennett
/ ; APPLICANT: Lex M. Cowsett
/ ; TITLE OF INVENTION: ANTISENSE MODULATION OF HER-2 EXPRESSION
/ ; FILE REFERENCE: RTS-0033
/ ; CURRENT APPLICATION NUMBER: US/09/663.834A
/ ; CURRENT FILING DATE: 2000-09-15
/ ; NUMBER OF SEQ ID NOS: 48
/ ; SEQ ID NO 10
/ ; LENGTH: 18
/ ; TYPE: DNA
/ ; ORGANISM: Artificial Sequence
/ ; FEATURE:
/ ; OTHER INFORMATION: Antisense Oligonucleotide
US-09-663-834A-10
Alignment Scores:
Pred. No.: 276 Length: 18
Score: 32.00 Matches: 6
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 32.00% Indels: 0
DB: 4 Gaps: 0
US-09-632-036F-1 (1-19) x US-09-663-834A-10 (1-18)
QY 10 ProAlaSerProGluThr 15
Db 18 CTGCCACGTCCGAGACC 1
RESULT 4
US-09-396-196G-18778
/ ; Sequence 18778, Application US/09396196G
/ ; Patent No. 6821724
/ ; GENERAL INFORMATION:
/ ; APPLICANT: Michael Mittmann
/ ; APPLICANT: David Mack
/ ; APPLICANT: David Lockhart
/ ; APPLICANT: Affymetrix, Inc.
/ ; TITLE OF INVENTION: Methods of Genetic Analysis
/ ; FILE REFERENCE: 3101.1
/ ; CURRENT APPLICATION NUMBER: US/09/396,196G
/ ; CURRENT FILING DATE: 1999-09-15
/ ; PRIOR APPLICATION NUMBER: 60/100,678
/ ; PRIOR FILING DATE: 1998-09-17
/ ; NUMBER OF SEQ ID NOS: 127806
/ ; SOFTWARE: FastSeq for Windows Version 4.0
/ ; SEQ ID NO 18778
/ ; LENGTH: 25
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 17, 2005, 22:52:13 ; Search time 206.738 Seconds
(without alignments)
544.046 Million cell updates/sec

Title: US-09-632-036F-1

Perfect score: 100

Sequence: 1 TGTDMKRLPASPETHLDM 19

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 3916100

Minimum DB seq length: 0
Maximum DB seq length: 40

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09632036/runat_16062005_151336_11008/app_query.fasta_1.1791
-DB=N_Geneseq_16Dec04 -Qfmt=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=40
-USPR=US09632036 @CGN 1.1 2129 @runat_16062005_151336_11008 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37	37.0	35	12 ADL72974	Hepatitis
2	36	36.0	29	10 ACH01230	Acho1230 Stem loop
3	36	36.0	29	10 ACH01236	Acho1236 Stem loop
C 4	35	35.0	19	11 ADL79101	Adl79101 Human HER
5	35	35.0	19	11 ADL78852	Adl78852 Human HER

C 6	35	35.0	30	2 AAQ27073	AAQ27073 HCV prime
7	35	35.0	32	12 ADN97675	Adn97675 Primer A.
C 8	35	35.0	37	9 ACC83681	Acc83681 Fragile X
C 9	34	34.0	33	10 ABZ25064	Abz25064 Complemen
10	34	34.0	39	2 AAV48328	Aav48328 Nucleotid
C 11	34	34.0	29	8 ACC42743	Acc42743 Human ULB
C 12	33	33.0	25	9 ACK30771	Ack30771 Human mic
C 13	33	33.0	25	9 ACK11533	Ack11533 Human mic
C 14	33	33.0	30	4 AAF83070	Aaf83070 TK promot
C 15	33	33.0	32	5 AAS04022	Aas04022 VNTR hydr
C 16	33	33.0	33	4 AAH75637	Aah75637 Human rib
C 17	33	33.0	33	6 ABA01842	Abao1842 Human lig
18	33	33.0	34	12 ACF36932	Acf36932 Human alp
19	33	33.0	34	12 ACF36928	Acf36928 Human alp
20	33	33.0	34	12 ACF36930	Acf36930 Human alp
21	33	33.0	34	12 ACF36934	Acf36934 Human alp
22	33	33.0	35	12 ADL72971	Adl72971 Hepatitis
C 23	33	33.0	37	9 ACC83732	Acc83732 Fragile X
24	32.5	32.5	36	2 AAQ58499	Aaq58499 Sequence
C 25	32	32.0	18	6 AAD38911	Aad38911 Human Her
26	32	32.0	18	13 ADQ75952	Adq75952 Rhesus mo
C 27	32	32.0	18	13 ADQ75948	Adq75948 Rhesus mo
28	32	32.0	19	11 ADL78850	Adl78850 Human HER
C 29	32	32.0	19	11 ADL79099	Adl79099 Human HER
C 30	32	32.0	27	3 AAZ59942	Aaz59942 Rat ribos
C 31	32	32.0	30	4 AAH21866	Aah21866 Mouse N-c
C 32	32	32.0	32	2 AAT18055	Aat18055 Sense pri
33	32	32.0	33	2 AAQ52044	Aaq52044 Breast ca
C 34	32	32.0	33	2 AAV72967	Aav72967 Human Doc
35	32	32.0	35	2 AAX03931	Aax03931 PCR prime
C 36	32	32.0	35	4 AAL46388	Aal46388 Rat GIRK
C 37	32	32.0	37	2 AAX91875	Aax91875 Porphorym
C 38	32	32.0	37	2 AAX54779	Aax54779 NK-kB ant
C 39	32	32.0	37	3 AAZ92747	Aaz92747 Saccharom
C 40	32	32.0	38	6 ABN84619	Abn84619 Anti-Stre
41	32	32.0	40	12 ADJ53680	Adj53680 HBV captu
42	31	31.0	21	6 ABX09462	Abx09462 Arteriosc
C 43	31	31.0	24	6 ABN84577	Abn84577 HSV thymi
44	31	31.0	25	4 AAH39091	Aah39091 SNP speci
C 45	31	31.0	25	9 ACI23138	AcI23138 Human mic

ALIGNMENTS

RESULT 1
ADL72974/c
ID ADL72974 standard; DNA; 35 BP.
XX
AC ADL72974;
XX
XX
DT 17-JUN-2004 (first entry)
XX
DE Hepatitis C virus E1E2 polyprotein plasmid PCR primer #6.
KW HCV; hepatitisvirus; vaccine; hepatitis; infection; plasmid; ss; primer;
KW PCR.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
PN WO2004024904-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-IB003882.
XX
PR 13-SEP-2002; 2002EP-00292254.
PR 03-MAR-2003; 2003EP-00290505.
XX
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX Bartosch B, Cosset F;
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 17, 2005, 22:54:33 ; Search time 1174.74 Seconds
(without alignments)
783.703 Million cell updates/sec

Title: US-09-632-036F-1

Perfect score: 100

Sequence: 1 TGTDMKRLPASPETHLDM 19

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 1692386

Minimum DB seq length: 0
Maximum DB seq length: 40

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO_spool/US09632036/runat_16062005_151337_11017/app_query.fasta_1.1791
-DB=GenEmbl -QPM=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCD=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=40
-USER=US09632036@cgn 1.1 9703 @runat_16062005_151337_11017 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_hrg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_ets.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	37	37.0	35	6	CQ785587 Sequence
C 2	37	37.0	35	6	CQ795329 Sequence
C 3	37	37.0	35	6	CQ867233 Sequence
C 4	35	35.0	30	9	X58166 Human T-cel

5	35	35.0	32	6	CQ801265 Sequence
C 6	33	33.0	30	6	AX107905 Sequence
C 7	33	33.0	32	6	AR154125 Sequence
8	33	33.0	35	6	CQ785584 Sequence
9	33	33.0	35	6	CQ795326 Sequence
10	33	33.0	35	6	CQ867230 Sequence
C 11	32	32.0	18	6	CQ841321 Sequence
12	32	32.0	18	6	CQ841325 Sequence
C 13	32	32.0	18	6	AR392095 Sequence
C 14	32	32.0	30	6	BD095106 N-Type ca
C 15	32	32.0	32	6	E10693 Primer, 9/1
16	32	32.0	33	6	I34919 Sequence 5
17	32	32.0	37	6	BD227235 Polyhydro
18	32	32.0	37	6	AR342651 Sequence
19	32	32.0	37	6	AR366501 Sequence
20	31	31.0	21	6	AX113664 Sequence
21	31	31.0	25	6	AX116764 Sequence
22	31	31.0	29	6	AX497614 Sequence
23	31	31.0	29	6	AX539926 Sequence
C 24	31	31.0	38	6	I58428 Sequence 12
25	31	31.0	38	6	AR200026 Sequence
26	31	31.0	38	6	AR227671 Sequence
27	31	31.0	39	10	MMIGH92
C 28	30	30.0	24	6	BD137624 Mutation
C 29	30	30.0	24	6	AR182613 Sequence
C 30	30	30.0	24	6	AR182615 Sequence
C 31	30	30.0	24	6	AR182617 Sequence
C 32	30	30.0	25	6	CQ619524 Sequence
C 33	30	30.0	25	6	CQ619525 Sequence
C 34	30	30.0	25	6	CQ620796 Sequence
C 35	30	30.0	25	6	CQ620797 Sequence
C 36	30	30.0	25	6	CQ620798 Sequence
C 37	30	30.0	25	6	CQ620799 Sequence
C 38	30	30.0	25	6	CQ620800 Sequence
C 39	30	30.0	25	6	AR460587 Sequence
C 40	30	30.0	25	6	AR460588 Sequence
C 41	30	30.0	25	6	AR461859 Sequence
C 42	30	30.0	25	6	AR461860 Sequence
C 43	30	30.0	25	6	AR461861 Sequence
C 44	30	30.0	25	6	AR461862 Sequence
C 45	30	30.0	25	6	AR461863 Sequence

ALIGNMENTS

RESULT 1	CQ785587/c	CQ785587	Sequence 7 from Patent EP1398371.	35 bp	DNA	linear	PAT 24-MAR-2004
LOCUS	CQ785587	CQ785587	GI:45720836				
DEFINITION	CQ785587	CQ785587	GI:45720836				
ACCESSION	CQ785587	CQ785587	GI:45720836				
VERSION	CQ785587.1	CQ785587.1	GI:45720836				
KEYWORDS	CQ785587.1	CQ785587.1	GI:45720836				
SOURCE	CQ785587.1	CQ785587.1	GI:45720836				
ORGANISM	CQ785587.1	CQ785587.1	GI:45720836				
REFERENCE	CQ785587.1	CQ785587.1	GI:45720836				
AUTHORS	CQ785587.1	CQ785587.1	GI:45720836				
TITLE	CQ785587.1	CQ785587.1	GI:45720836				
JOURNAL	CQ785587.1	CQ785587.1	GI:45720836				
FEATURES	CQ785587.1	CQ785587.1	GI:45720836				
source	CQ785587.1	CQ785587.1	GI:45720836				
Location/Qualifiers	CQ785587.1	CQ785587.1	GI:45720836				
1..35	CQ785587.1	CQ785587.1	GI:45720836				
/organism="synthetic construct"	CQ785587.1	CQ785587.1	GI:45720836				
/mol_type="unassigned DNA"	CQ785587.1	CQ785587.1	GI:45720836				
/db_xref="taxon:32630"	CQ785587.1	CQ785587.1	GI:45720836				
/note="Description of Artificial Sequence: primer"	CQ785587.1	CQ785587.1	GI:45720836				

ORIGIN	CQ785587/c	CQ785587	Sequence 7 from Patent EP1398371.	35 bp	DNA	linear	PAT 24-MAR-2004
Alignment Scores:	CQ785587/c	CQ785587	Sequence 7 from Patent EP1398371.	35 bp	DNA	linear	PAT 24-MAR-2004
Pred. No.:	CQ785587/c	CQ785587	Sequence 7 from Patent EP1398371.	35 bp	DNA	linear	PAT 24-MAR-2004
Score:	CQ785587/c	CQ785587	Sequence 7 from Patent EP1398371.	35 bp	DNA	linear	PAT 24-MAR-2004

